

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



**PENERAPAN ALGORITME RABIN-KARP DAN COSINE
SIMILARITY UNTUK PEMERIKSAAN KESAMAAN
DOKUMEN TUGAS MAKALAH MAHASISWA
(STUDI KASUS: TEKNIK INFORMATIKA UIN SUSKA RIAU)**

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh:

ELVY MARISHA
11551205100



FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU

2019

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© H

LEMBAR PERSETUJUAN
PENERAPAN ALGORITME *RABIN-KARP* DAN *COSINE*
***SIMILARITY* UNTUK PEMERIKSAAN KESAMAAN**
DOKUMEN TUGAS MAKALAH MAHASISWA
(STUDI KASUS: TEKNIK INFORMATIKA UIN SUSKA RIAU)


TUGAS AKHIR

Oleh:


ELVY MARISHA
11551205100

Telah diperiksa dan disetujui sebagai laporan Tugas Akhir
di Pekanbaru, pada tanggal 25 November 2019

Pembimbing I,


Suwanto Sanjaya, ST., M.Kom
NIK. 130 517 103

Pembimbing II,


Elvia Budianita, ST., M.Cs
NIP. 19860629 201503 2 007

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PENGESAHAN

PENERAPAN ALGORITME *RABIN-KARP* DAN *COSINE*
SIMILARITY UNTUK PEMERIKSAAN KESAMAAN
DOKUMEN TUGAS MAKALAH MAHASISWA
(STUDI KASUS: TEKNIK INFORMATIKA UIN SUSKA RIAU)

TUGAS AKHIR

Oleh:

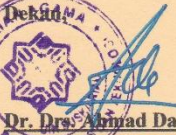
ELVY MARISHA
11551205100

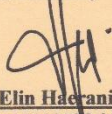
Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar sarjana Teknik Informatika
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 25 November 2019

Pekanbaru, 25 November 2019

Mengesahkan

Ketua Jurusan,


Dr. Drs. Ahmad Darmawi, M.Ag.
NIP. 19660604 199203 1 004


Dr. Elin Haerani, S.T., M.Kom.
NIP. 19810523 200710 2 003

DEWAN PENGUJI

Ketua	: Muhammad Fikry, S.T., M.Sc.
Sekretaris	: Suwanto Sanjaya, S.T., M.Kom.
Pembimbing II	: Elvia Budianita, S.T., M.Cs.
Penguji I	: Dr. Elin Haerani, S.T., M.Kom.
Penguji II	: Pizaini, S.T., M.Kom.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman, dan tanggal pinjam.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Pekanbaru, 25 November 2019

Yang membuat pernyataan,

ELVY MARISHA
11551205100

UIN SUSKA RIAU

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

LEMBAR PERSEMBAHAN



Alhamdulillahirobbil'aalamiin...

Sembah sujud dan syukur kepada Allah . Taburan cinta dan kasih sayang Mu telah memberikan ku kekuatan, membekaliku dengan ilmu serta memperkenalkanku dengan cinta. Atas rahmat dan karunia yang telah Engkau berikan, Alhamdulillah Tugas Akhir ini dapat terselesaikan. Sholawat dan salam selalu terlimpahkan kepada rasul kita Muhammad ﷺ.

Kupersembahkan karya ini kepada orang yang sangat kukasihi dan kusayangi.

Mama dan Papa Tercinta

Kepada mama dan papa tercinta. Sebagai tanda bukti, hormat dan rasa terima kasih yang tiada terhingga. Kupersembahkan kepada Mama (Zulheda) dan Papa (Masril) yang telah memberikan doa dan kasih sayang dari lahir hingga dewasa. Semoga hasil karya ini menjadi salah satu kebahagiaan yang bisa kuberikan. Terima kasih, Ma... Terima Kasih, Pa...

Abang dan Adik Tersayang

Sebagai tanda terima kasih. Kupersembahkan karya ini untuk abang Putra Idaman, Adik Ivana Ramadhani dan semua keluarga terdekatku. Terima kasih telah memberikan doa, motivasi dan semangat sehingga terselesaikan Tugas Akhir ini. Semoga ini menjadi hal terbaik yang bisa kuberikan. Terima kasih...

Teman-teman

Buat teman-temanku yang telah menemani perjuanganku dalam menyelesaikan Tugas Akhir ini. Terima kasih ku ucapkan untuk teman-teman seperjuangan. Sekali lagi Terima kasih teman-temanku...

Dosen Pembimbing Tugas Akhir

Sebagai tanda terima kasih. Saya ucapkan terima kasih banyak kepada Bapak Suwanto Sanjaya, ST., M.Kom dan Ibu Elvia Budianita, ST., M.Cs yang telah memberikan arahan dan bimbingan. Terima kasih bapak dan ibu atas semua ilmu, arahan, nasehat dan candaan yang telah bapak dan ibu berikan. Semoga Allah selalu melindungi langkah bapak dan ibu dimanapun bapak dan ibu berada....

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

**PENERAPAN ALGORITME RABIN-KARP DAN COSINE
SIMILARITY UNTUK PEMERIKSAAN KESAMAAN
DOKUMEN TUGAS MAKALAH MAHASISWA
(STUDI KASUS: TEKNIK INFORMATIKA UIN SUSKA RIAU)**

ELVY MARISHA
11551205100

Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau

ABSTRAK

Kesamaan dokumen (*document similarity*) merupakan fondasi dari sistem kecerdasan dalam pemrosesan data seperti *information retrieval* dan klasifikasi teks. Berdasarkan hasil wawancara yang dilakukan terhadap beberapa dosen Teknik Informatika UIN SUSKA Riau, didapatkan bahwa masih banyak ditemukan kasus kesamaan dokumen tugas makalah antara satu mahasiswa dengan mahasiswa lain, sehingga perlu dilakukan proses pemeriksaan kesamaan dokumen tugas makalah mahasiswa. Pada penelitian ini, algoritme yang diterapkan untuk melakukan proses pemeriksaan kesamaan dokumen adalah algoritme *Rabin-Karp* dan *Cosine Similarity*. Algoritme *Rabin-Karp* digunakan untuk proses *preprocessing* dan ekstraksi nilai *hash*. Sedangkan *Cosine similarity* digunakan untuk menghitung nilai persentase kesamaan dari dokumen yang diuji. Hasil pengujian tata letak kalimat didapatkan bahwa nilai *similarity* akan tetap sama walaupun letak kalimat telah diubah. Selanjutnya, hasil pengujian dokumen yang sama akan menghasilkan persentase *similarity* 100%. Sebaliknya, hasil pengujian dokumen yang tidak sama akan menghasilkan persentase *similarity* 0%. Selain itu pengujian dua dokumen dengan nilai *K* yang digunakan 3, 5, 6, dan 7 diperoleh hasil bahwa semakin kecil nilai *K-gram* maka semakin tinggi hasil *similarity* yang diperoleh dan nilai *similarity* tertinggi terletak pada *K*= 3 yaitu sebesar 18,54 %. Kemudian, hasil pengujian perbandingan antara sistem pemeriksaan kesamaan dokumen dengan *plagiarism checker* x dari 15 *file* dokumen, diperoleh nilai *similarity* tertinggi 10,29 % pada sistem pemeriksaan dokumen dan pada *plagiarism checker* x nilai *similarity* tertinggi 14,06 %.

Kata Kunci : Dokumen, Pemeriksaan Kesamaan, *Rabin-Karp*, *Cosine similarity*, *K-gram*, Tugas Makalah

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

IMPLEMENTATION OF RABIN-KARP ALGORITHM AND COSINE SIMILARITY FOR SIMILARITY EXAMINATION OF STUDENTS' PAPERS ASSIGNMENT DOCUMENTS (CASE STUDY: TEKNIK INFORMATIKA UIN SUSKA RIAU)

ELVY MARISHA
11551205100

*Informatics Engineering
Faculty of Science and Technology
State Islamic University of Sultan Syarif Kasim Riau*

ABSTRACT

Document similarity is the foundation of intelligence systems in data processing such as information retrieval and text classification. Based on the results of interviews conducted with several Informatics Engineering lecturers of UIN SUSKA Riau, it was found that there were still many cases of similarity of paper assignment documents between one student and another student, so it was necessary to check the similarity of student paper assignment documents. In this study, the algorithm applied to check the document similarity is the Rabin-Karp algorithm and Cosine Similarity. The Rabin-Karp algorithm is used for the process of preprocessing and extracting hash values. While Cosine similarity is used to calculate the percentage value of similarity of the document being tested. The results of the sentence layout test found that the similarity value will remain the same even though the sentence location has been changed. Furthermore, the results of testing the same document will produce a percentage of 100% similarity. Conversely, the results of testing documents that are not the same will produce a percentage of 0% similarity. Besides testing two documents with K values used 3, 5, 6, and 7 obtained the results that the smaller the K-gram value, the higher the similarity results obtained and the highest similarity value lies at K = 3 which is equal to 18.54%. Then, the comparison test results between the document similarity checking system with the plagiarism checker x from 15 document files, obtained the highest similarity value of 10.29% in the document inspection system and the plagiarism checker x the highest similarity value of 14.06%.

Keywords : Documents, Similarity Checking, Rabin-Karp, Cosine similarity, K-gram, Paper Assignment

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji syukur penulis ucapkan kepada Allah SWT, karena rahmat dan karunia-Nya penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Penerapan Algoritme Rabin-Karp dan Cosine Similarity untuk Pemeriksaan Kesamaan Dokumen Tugas Makalah Mahasiswa”**. Shalawat beriring salam penulis haturkan kepada junjungan besar Nabi Muhammad SAW yang telah membawa kita dari zaman jahiliah menuju zaman yang penuh ilmu pengetahuan.

Tugas Akhir ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana strata satu pada Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Selama menyelesaikan Tugas Akhir ini, penulis mendapatkan banyak pengetahuan, bimbingan, arahan, dukungan serta doa dari berbagai pihak yang telah bersedia meluangkan waktu dan pikirannya baik materil maupun moril. Sehingga pada kesempatan ini penulis ingin mengucapkan terimakasih kepada:

1. Bapak Prof. Dr. H. Akhmad Mujahidin, S.Ag., M.Ag selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Drs. Ahmad Darmawi, M.Ag selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Ibu Dr. Elin Haerani, ST., M.Kom selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Bapak Suwanto Sanjaya, ST., M.Kom selaku dosen pembimbing I Tugas Akhir yang telah banyak membantu penulis dan memberikan bimbingan, arahan, nasihat, serta waktu sehingga penulis mampu menyelesaikan Tugas Akhir ini.
5. Ibu Elvia Budianita, ST., M.Cs selaku dosen pembimbing akademik dan dosen pembimbing II Tugas Akhir yang telah memberikan banyak pengetahuan, bimbingan, arahan, nasehat, serta waktu sehingga penulis mampu menyelesaikan Tugas Akhir ini dengan sebaik mungkin.
6. Ibu Dr. Elin Haerani, ST., M.Kom selaku dosen penguji I yang telah banyak membantu serta memberi saran dalam penyempurnaan Tugas Akhir ini.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

7. Bapak Pizaini, ST., M.Kom selaku dosen penguji II yang telah banyak membantu serta memberi saran dalam penyempurnaan Tugas Akhir.
8. Seluruh Bapak dan Ibu dosen Teknik Informatika UIN Suska Riau yang telah memberikan ilmu dan pengetahuan selama masa perkuliahan.
9. Ayahanda Masril dan Ibunda Zulheda yang tidak pernah letih untuk selalu memberikan semangat, nasihat, dan kasih sayang yang tidak pernah pudar. Selalu menjadi tempat berkeluh kesah semua permasalahan penulis. Semoga Allah *subhanahu wa ta'ala* selalu memberikan kesehatan dan perlindungan dimanapun Ayahanda dan Ibunda berada.
10. Kakanda Putra Idaman dan Adinda Ivana Ramadhani yang selalu menghibur penulis disaat penulis merasa jenuh dengan dunia perkuliahan dan memberikan penulis semangat untuk selalu berjuang. Terima kasih, bang dan pana.
11. Keluarga besar penulis yang telah menginspirasi, memberikan semangat, dan dukungan kepada penulis.
12. Teman-teman seperjuangan kelas C angkatan 2015 yang selalu berjuang bersama dan selalu saling mendukung. Teristimewa kepada, Indah Puji Lestari Herda, Miftahul Khairat, Ika Yolanda, dan Wilda Anisah yang selalu membantu penulis dalam segala keadaan.
13. Teman-teman senasib yang tersayang dan terkasih kepada Nindi Henisa, Nanda Hardiyanti, Rada Guspita Wanda, dan Shintiya Monalisa terimakasih karena selalu ada disisi penulis dalam keadaan baik senang maupun susah.
14. Semua pihak yang tidak dapat penulis sebutkan satu persatu. Terimakasih atas dukungan baik moril maupun materil dalam pengerjaan Tugas Akhir ini.

Semoga laporan ini dapat bermanfaat bagi penulis khususnya maupun pembaca umumnya. Penulis sadar masih banyak kekurangan oleh karena itu penulis berharap bisa mendapatkan masukan dari pembaca atas isi laporan ini. Akhir kata penulis ucapkan terima kasih dan selamat membaca.

Pekanbaru, 25 November 2019
Penulis

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR ISI

LEMBAR PERSETUJUAN	i
LEMBAR PENGESAHAN	ii
LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL.....	iii
LEMBAR PERNYATAAN	iv
LEMBAR PERSEMBAHAN	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xvi
DAFTAR LAMPIRAN	xviii
DAFTAR SIMBOL	xix
BAB I PENDAHULUAN.....	I-1
1.1 Latar Belakang	I-1
1.2 Rumusan Masalah	I-5
1.3 Batasan Masalah.....	I-5
1.4 Tujuan Penelitian.....	I-5
1.5 Sistematika Penulisan.....	I-6
BAB II LANDASAN TEORI	II-1
2.1 Kesamaan Dokumen	II-1
2.2 Metode Pengukuran Kesamaan Dokumen	II-1
2.3 <i>Text Mining</i>	II-2
2.4 <i>Text Preprocessing</i>	II-2
2.5 <i>Rabin-Karp</i>	II-3
2.5.1 <i>Case folding</i>	II-3
2.5.2 <i>Filtering</i>	II-4
2.5.3 Menentukan <i>K-gram</i>	II-4
2.5.4 <i>Hashing</i>	II-5
2.5.5 Membandingkan Nilai <i>Hash</i> yang Sama.....	II-7
2.6 <i>Similarity</i>	II-9

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.7 Penelitian Terkait	II-10
------------------------------	-------

BAB III METODOLOGI PENELITIAN III-1

3.1 Studi Pustaka	III-1
3.2 Perumusan Masalah.....	III-2
3.3 Pengumpulan Data	III-2
3.4 Analisa.....	III-2
3.4.1 Analisa Kebutuhan Data.....	III-2
3.4.2 Analisa Tahapan Ekstraksi Nilai Hash Menggunakan Algoritme Rabin-Karp	III-3
3.4.3 Analisa Fungsional Sistem	III-4
3.5 Perancangan Aplikasi.....	III-5
3.6 Implementasi	III-6
3.7 Pengujian.....	III-6
3.8 Kesimpulan dan saran	III-7

BAB IV ANALISA DAN PERANCANGANIV-1

4.1 Analisa.....	IV-1
4.1.1 Analisa Kebutuhan Data.....	IV-1
4.1.2 Analisa Proses Ekstraksi Nilai <i>Hashing</i> Menggunakan Algoritme <i>Rabin-Karp</i> Pada Dokumen 1	IV-1
4.1.3 Proses Ekstraksi Nilai <i>Hashing</i> Dokumen 2	IV-14
4.1.4 Proses Pencarian Nilai <i>Hash</i> yang Sama.....	IV-21
4.1.5 Proses Perhitungan Persentase Kesamaan (<i>Similarity</i>)	IV-22
4.1.6 Analisa Fungsional Sistem	IV-24
4.1.6.1 <i>Use case Diagram</i>	IV-24
4.1.6.2 <i>Use case Spesification</i>	IV-25
4.1.6.3 <i>Sequence Diagram</i>	IV-26
4.1.6.4 <i>Class Diagram</i>	IV-29
4.2 Perancangan	IV-30
4.2.1 Perancangan Struktur Menu	IV-30
4.2.2 Perancangan Tampilan Antarmuka	IV-30

BAB V IMPLEMENTASI DAN PENGUJIAN..... V-1

5.1 Implementasi	V-1
------------------------	-----

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

5.1.1	Batasan Implementasi.....	V-1
5.1.2	Lingkungan Implementasi	V-1
5.1.3	Hasil Implementasi <i>Interface</i> Sistem	V-2
5.2	Pengujian.....	V-8
5.2.1	Pengujian Fungsional Algoritme.....	V-8
5.2.2	Pengujian Tata Letak Kalimat.....	V-22
5.2.3	Pengujian Dokumen yang Sama.....	V-25
5.2.4	Pengujian Dokumen yang Tidak Memiliki Kesamaan.....	V-26
5.2.5	Pengujian Parameter <i>K-gram</i>	V-30
5.2.6	Pengujian Banyak Dokumen	V-37
5.2.7	Pengujian Perbandingan Sistem Pemeriksaan Kesamaan Dokumen Dengan Aplikasi Lain	V-41
5.2.8	Kesimpulan Pengujian.....	V-45
BAB VI PENUTUP		VI-1
6.1	Kesimpulan.....	VI-1
6.2	Saran.....	VI-2
DAFTAR PUSTAKA		xxii
LAMPIRAN A		A-1
LAMPIRAN B		B-1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR GAMBAR

Gambar	Halaman
2.1 Fingerprint awal	II-7
2.2 Pergeseran <i>Fingerprint</i>	II-8
2.3 Perbandingan Kedua	II-8
2.4 perbandingan keempat (nilai <i>hash</i> yang sama)	II-8
2.5 perbandingan kelima (<i>string</i> ditemukan)	II-9
3.1 Tahapan Penelitian	III-1
4.1 <i>Flowchart</i> algoritme <i>Rabin-Karp</i>	IV-2
4.2 <i>Flowchart</i> Tahapan <i>Preprocessing</i>	IV-4
4.3 <i>Flowchart</i> Proses <i>Case folding</i>	IV-4
4.4 <i>Flowchart</i> Proses <i>Filtering</i>	IV-6
4.5 <i>Flowchart</i> Pembentukan <i>K-gram</i>	IV-7
4.6 <i>Flowchart</i> Proses <i>Hashing</i>	IV-10
4.7 <i>Flowchart</i> Proses Pemeriksaan Kesamaan	IV-22
4.8 <i>Use case Diagram</i>	IV-24
4.9 <i>Sequence Diagram</i> Melakukan Cek <i>Similarity</i> Dua Dokumen	IV-27
4.10 <i>Sequence Diagram</i> Melakukan Cek <i>Similarity</i> Banyak Dokumen	IV-28
4.11 <i>Class Diagram</i>	IV-29
4.12 Struktur Menu Sistem	IV-30
4.13 Tampilan Menu <i>Home</i>	IV-31
4.14 Tampilan Menu Pemeriksaan Dua Dokumen	IV-32
4.15 Tampilan Tab <i>Case folding</i>	IV-32
4.16 Tampilan Tab <i>Tokenizing</i>	IV-33
4.17 Tampilan Tab <i>Hashing</i>	IV-33
4.18 Tampilan Tab <i>Similarity</i>	IV-34
4.19 Tampilan Tab <i>Result</i>	IV-34
4.20 Tampilan Halaman Pemeriksaan Banyak Dokumen	IV-35
4.21 Tampilan Halaman <i>Result</i> Pemeriksaan Banyak Dokumen	IV-36
5.1 Tampilan Halaman <i>Home</i>	V-2
5.2 Tampilan Halaman Pemeriksaan Dua Dokumen	V-3
5.3 Tampilan Tab Isi Dokumen	V-3

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

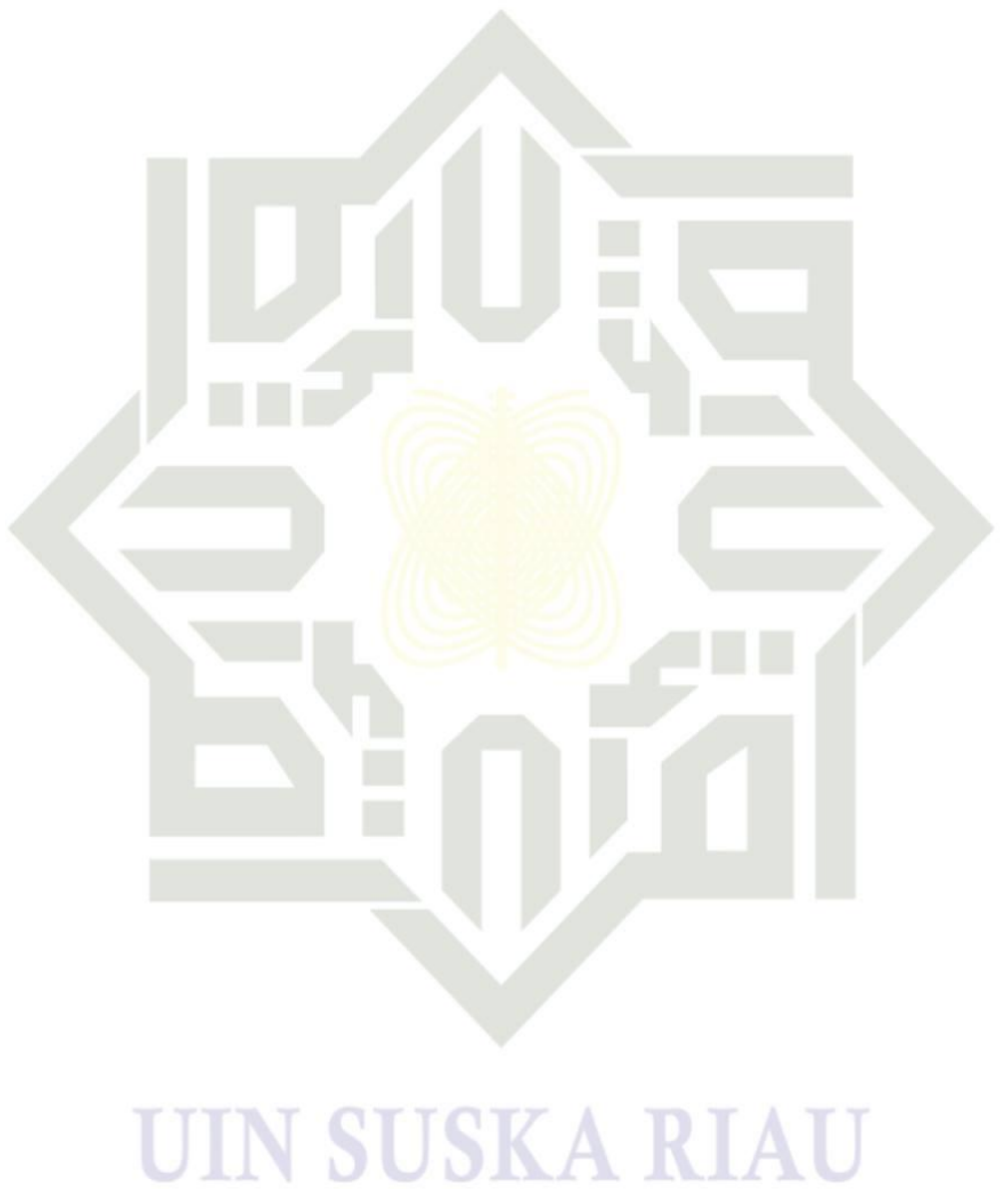
© Hak Cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

5.4 Tampilan Tab <i>Case folding & Stoplist Removal</i>	V-4
5.5 Tampilan Tab <i>Tokenization</i>	V-4
5.6 Tampilan Tab <i>Hashing</i>	V-5
5.7 Tampilan Tab <i>Similarity</i>	V-5
5.8 Tampilan Tab <i>Result</i>	V-6
5.9 Tampilan Halaman Pemeriksaan Banyak Dokumen.....	V-6
5.10 Tampilan Halaman Pemeriksaan Setelah diinputkan Dokumen.....	V-7
5.11 Tampilan Halaman <i>Result</i>	V-7
5.12 Hasil <i>Case folding & Filtering (Stoplist Removal)</i>	V-10
5.13 Hasil Pengujian <i>Tokenization & Hashing</i>	V-19
5.14 Hasil Pengujian <i>Similarity</i>	V-21
5.15 Hasil Pengujian <i>Highlight</i> Kedua Dokumen.....	V-22
5.16 Hasil <i>similarity</i> Dokumen 1 Asli dan yang telah Diubah	V-24
5.17 <i>Highlight</i> Dokumen Asli	V-25
5.18 <i>highlight</i> dokumen yang diubah.....	V-25
5.19 Hasil Pengujian Dokumen yang Sama	V-26
5.20 <i>Highlight</i> Pengujian Dokumen Yang Sama	V-26
5.21 Hasil <i>Similarity</i> Dokumen Dengan Topik Sama Tata Bahasa Berbeda....	V-28
5.22 <i>Highlight</i> Dokumen Dengan Topik Sama Tata Bahasa Berbeda.....	V-28
5.23 <i>Similarity</i> Dokumen Dengan Topik dan Tata Bahasa Berbeda	V-30
5.24 <i>Highlight</i> Dokumen Dengan Topik dan Tata Bahasa Berbeda.....	V-30
5.25 Hasil Pengujian 1	V-31
5.26 Hasil Pengujian 2	V-33
5.27 Hasil Pengujian 3	V-34
5.28 Hasil Pengujian 4	V-35
5.29 Hasil Pengujian 5	V-37
5.30 Hasil Pengujian Banyak Dokumen	V-40
5.31 Hasil Pengujian Sistem Pemeriksaan Dokumen	V-42
5.32 Hasil Pengujian Aplikasi <i>Plagiarism Checker X</i>	V-42
5.33 Hasil Pengujian Banyak Dokumen Sistem Pemeriksaan Dokumen	V-44
5.34 Hasil Pengujian Banyak Dokumen Aplikasi <i>Plagiarism Checker X</i>	V-44
5.35 Hasil Pengujian Banyak Dokumen Aplikasi <i>Plagiarism Checker X</i>	V-45

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

5.36 Grafik Perbandingan Sistem Pemeriksaan Kesamaan Dokumen dengan Plagiarism Checker X..... V-50



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

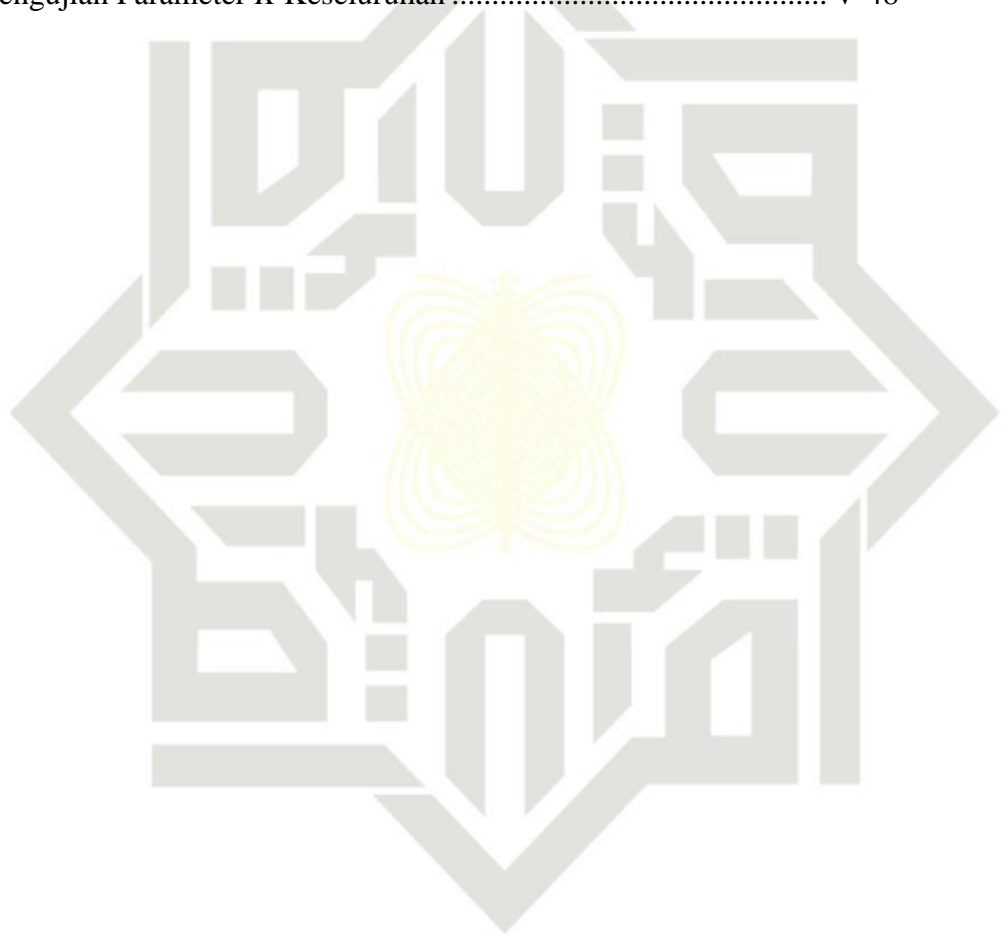
DAFTAR TABEL

Tabel	Halaman
2.1 Tabel ASCII	II-6
2.2 Penelitian Terkait	II-10
4.1 Contoh Dokumen 1	IV-3
4.2 Dokumen 1 Setelah Melalui Proses <i>Case folding</i>	IV-5
4.3 Dokumen 1 Setelah melalui Proses <i>Filtering (Stoplist Removal)</i>	IV-6
4.4 Hasil <i>K-gram</i> dokumen 1	IV-7
4.5 Nilai <i>hash</i> setiap <i>K-gram</i> dokumen 1	IV-12
4.6 Contoh Dokumen 2	IV-15
4.7 Dokumen 2 Setelah Melalui Proses <i>Case Folding</i>	IV-15
4.8 Dokumen 2 Setelah Melalui Proses <i>Filtering (Stoplist Removal)</i>	IV-16
4.9 Hasil <i>K-Gram</i> Dokumen 2	IV-16
4.10 Nilai <i>Hash</i> Setiap <i>K-Gram</i> Dokumen 2	IV-19
4.11 <i>Hashing</i> yang Sama Pada Dokumen 1 Dan Dokumen 2	IV-21
4.12 Hasil <i>Highlight</i> Kata Yang Sama Kedua Dokumen	IV-23
4.13 <i>Use Case Spesification</i> Melakukan Cek <i>Similarity</i> Dua Dokumen	IV-25
4.14 <i>Use Case Spesification</i> Melakukan Cek <i>similarity</i> Banyak Dokumen	IV-26
5.1 Hasil Pengujian <i>Case folding & Filtering (Stoplist Removal)</i>	V-8
5.2 Hasil Pengujian <i>Tokenization</i> dan <i>hashing</i> Dokumen 1	V-10
5.3 Hasil Pengujian <i>Tokenization</i> dan <i>hashing</i> Dokumen 2	V-16
5.4 Hasil Pengujian <i>Similarity</i>	V-20
5.5 Isi Dokumen 1 Asli dan yang Telah Diubah	V-22
5.6 Isi Dokumen 2 Asli dan yang Telah Diubah	V-23
5.7 Hasil <i>Similarity</i>	V-24
5.8 Pengujian Dokumen yang Sama	V-25
5.9 Perbandingan Isi Dokumen 1	V-27
5.10 Perbandingan Isi Dokumen 2	V-27
5.11 Perbandingan Isi Dokumen yang Diuji	V-29
5.12 Hasil Pengujian 1	V-31
5.13 Hasil Pengujian 2	V-32

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

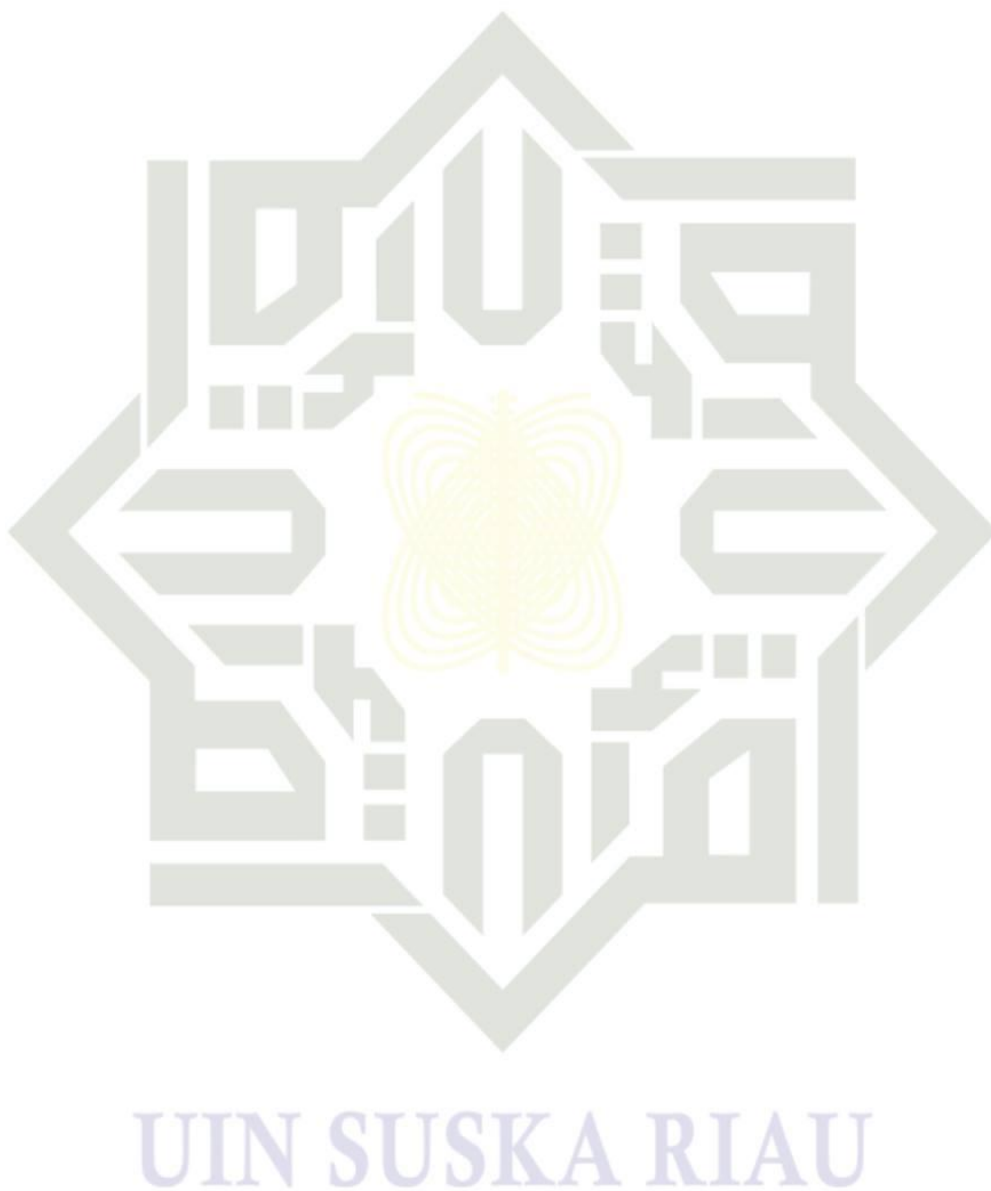
5.14 Hasil Pengujian 3	V-33
5.15 Hasil Pengujian 4	V-35
5.16 Hasil Pengujian 5	V-36
5.17 Pengujian Banyak Dokumen.....	V-37
5.18 Perbandingan Hasil <i>Similarity</i> Dua Dokumen	V-41
5.19 Hasil <i>Similarity</i> Menggunakan Sistem Pemeriksaan Dokumen	V-43
5.20 Hasil <i>Similarity</i> Menggunakan Aplikasi <i>Plagiarsm Checker X</i>	V-43
5.21 Hasil Pengujian Parameter <i>K</i> Keseluruhan	V-48



UIN SUSKA RIAU

DAFTAR LAMPIRAN

Lampiran	Halaman
A BUKTI WAWANCARA	A-1
B TABEL HASIL PERHITUNGAN NILAI <i>HASHING</i>	B-1



Hak Cipta Dilindungi Undang-Undang

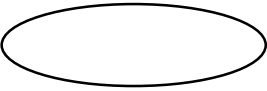
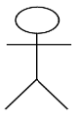

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang


1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR SIMBOL

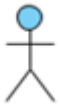



Simbol Use case Diagram

	<i>Use case</i> Menggambarkan interaksi antara aktor dengan sistem
	<i>Aktor</i> Menggambarkan pengguna sistem, yang berinteraksi dengan sistem lain untuk berbagi, mengirim, dan menerima informasi
	<i>Association</i> Penghubung antara objek satu dan objek lainnya.

Simbol Class Diagram

<div style="border: 1px solid black; padding: 5px; width: fit-content;"> <p>Class Name</p> <hr/> <p>+ attribute1:type = defaultValue + attribute2:type - attribute3:type</p> <hr/> <p>+ operation1(params):returnType - operation2(params) - operation3()</p> </div>	<i>Class</i> Himpunan dari objek yang berbagi atribut dan operasi yang sama.
	<i>Association</i> Relasi antar kelas dengan makna umum dan biasanya diikuti <i>multiplicity</i> .

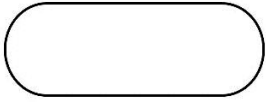
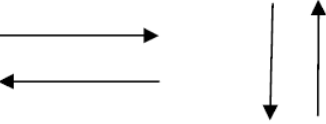

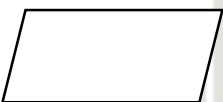
Simbol Sequence Diagram

	<i>Actor</i> Menggambarkan orang yang berinteraksi dengan sistem.
	<i>Life line</i> Menggambarkan sebuah peran atau instace object yang berpartisipasi dalam interaksi-interaksi yang di modelkan.
	<i>Object Message</i> Menggambarkan pesan/hubungan antar objek yang menunjukkan urutan kejadian yang terjadi.
	<i>Message to self</i>

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Simbol Flowchart

	Sebuah pesan yang menggambarkan eksekusi atau panggilan operasi didalam <i>lifeline</i> objek yang sama.
	Terminator Simbol permulaan (<i>start</i>) atau akhir (<i>end</i>) dari suatu kegiatan.
	Flow Direction Menghubungkan simbol yang satu dengan simbol yang lain.
	Process Menyatakan suatu tindakan (proses) baik yang dilakukan oleh <i>user</i> atau komputer.
	Input/Output Menyatakan proses <i>input</i> atau <i>output</i> tanpa tergantung jenis peralatan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB I PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang terus meningkat membuat informasi yang disediakan tidak lagi berbentuk *hardcopy* yang berupa berkas-berkas dokumen atau buku-buku yang disimpan pada rak-rak atau lemari buku. Namun dokumen, buku, ataupun media informasi lainnya sekarang sudah tersimpan dalam bentuk *softfile* atau dokumen digital. Dokumen-dokumen yang berbentuk digital tersebut dapat dengan mudah diduplikasi, tetapi hal tersebut banyak disalahgunakan oleh pengguna teknologi informasi dengan melakukan tindakan manipulasi pada *file* atau dokumen-dokumen yang bersumber dari pengguna lain (Wicaksono & Suyanto, 2012).

Pada dunia pendidikan, khususnya pada tingkat perguruan tinggi, kesamaan dokumen antar mahasiswa bukan suatu hal yang dianggap baru, namun sudah terjadi semenjak dahulu sebelum teknologi informasi masih belum secanggih seperti sekarang ini. Pada penelitian yang dilakukan oleh (Pradiansyah, 2015) sebanyak 120 responden yang merupakan mahasiswa dan mahasiswi di 12 fakultas yang diambil secara *representative* didapat sebanyak 85,5% mahasiswa dan mahasiswi ketika mengerjakan tugas-tugas kuliah tidak mencantumkan sumber atau referensi pada tugas yang mereka kerjakan, hal ini menyebabkan banyaknya terdapat dokumen tugas mahasiswa yang sama satu sama lain. Pelajar maupun mahasiswa banyak melakukan tindakan manipulasi pada pembuatan tugas seperti makalah, karya ilmiah, maupun tugas akhir atau skripsi, hal tersebut berdampak terhadap kreatifitas seseorang yang tidak tersalurkan karena lebih cenderung untuk meniru kreatifitas dari orang lain (Ernawati, Anindito, & Atmojo, 2014).

Pada penelitian ini, dilakukan wawancara terhadap beberapa dosen Teknik Informatika UIN SUSKA Riau, dari beberapa dosen menyatakan bahwa masih ditemukannya kasus kesamaan dokumen antar mahasiswa terhadap tugas-tugas yang diberikan, hal ini dapat berdampak pada rusaknya kualitas belajar dari mahasiswa. Bentuk kesamaan yang ditemukan beragam, baik *copy-paste* sebagian, *copy-paste* secara keseluruhan maupun *copy-paste* sebagian dan parafrase sebagian.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Pemeriksaan antar dokumen tersebut masih dilakukan dengan cara membaca satu persatu dokumen tugas mahasiswa dan untuk pemeriksaan kesamaan dapat dilihat dari penyusunan bab, jenis *font*, ataupun penyusunan kalimat pada dokumen. Hal tersebut tentu membutuhkan waktu untuk melakukan pencarian terhadap dokumen-dokumen yang bersangkutan.

Beberapa penelitian sebelumnya telah dilakukan mengenai pemeriksaan kesamaan kata pada teks dokumen, salah satunya penelitian yang dilakukan oleh (Mahmood, Qadir, & Afzal, 2013) dengan menggunakan algoritme *Semantic Social Network Analysis* pada grafik kutipan RDF, penelitian ini berfokus pada dokumen yang berisi referensi kutipan ke dokumen lain, karena set data yang digunakan pada teknik ini berisi kutipan *link* antara dokumen. Jika beberapa dokumen tidak mengandung kutipan referensi, maka teknik yang digunakan pada penelitian ini tidak dapat memberikan hasil kesamaan antar dokumen. Hasil *similarity* yang diperoleh dari grafik kutipan RDF 72,33% langkah pada teknik ini hampir sama dengan pola yang diidentifikasi manusia secara manual, tetapi perhitungan *similarity* secara manual memiliki hasil yang lebih baik dari pada analisis grafik kutipan RDF karena heuristik dan kecerdasan dari manusia dalam menganalisis kutipan *link*.

Pada penelitian yang dilakukan oleh (Wibowo, Sudarmadi, & Barmawi, 2013) dalam pendeteksian plagiarisme dokumen berbahasa indonesia, penelitian ini menggunakan kinerja algoritme *fingerprint* dan *winnowing*, hasil tertinggi pada analisis tes akurasi algoritme *fingerprint* sebesar 92,8% dengan nilai *threshold* 0,1 dan nilai N-gram berada di 3, algoritme *winnowing* sebesar 91,8%. hasil yang berbeda ditunjukkan pada tingkat relevansi akurasi dengan topik, hasil akurasi dari algoritme *winnowing* mendapatkan korelasi yang lebih kuat 37,1% dibandingkan *fingerprint* 33,6%.

Pada algoritme yang berbeda yaitu dengan menggunakan algoritme *Latent Semantic Analysis* (LSA) dan *Vector Space Model* (VSM), penelitian yang dilakukan oleh (Yu, Chen, & Shi, 2017) melakukan perbandingan algoritme LSA dan VSM, dari penelitian ini perbandingan akurasi dengan menerapkan protokol dari VSM (diberi label “VSM”) dan protokol LSA dengan *rank* 100 (diberi label “LSA-100”), *rank* 150 (diberi label “LSA-150”), dan *rank* 200 (diberi label “LSA-

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

200”) dimana akurasi terbaik didapatkan dari LSA-150 dengan akurasi 0,86 dan VSM dengan 0,80. waktu perhitungan yang diperlukan oleh protokol VSM yaitu 103 menit sementara LSA-100 membutuhkan 72 menit. Tetapi pada penelitian ini menyimpulkan meski protokol LSA lebih cepat dari protokol VSM, LSA masih tidak terlalu efisien untuk menangani perhitungan kesamaan antar kumpulan dokumen yang berskala besar.

Berkaitan dengan judul penelitian yang akan diteliti mengenai algoritme *Rabin-Karp*, terdapat beberapa penelitian terkait diantaranya, penelitian yang dilakukan oleh (Leonardo & Hansun, 2017) melakukan perbandingan antara algoritme *Rabin-Karp* dengan algoritme *Jaro-Winkler*. Berdasarkan hasil analisis dari eksperimen, secara umum algoritme *Rabin-Karp* lebih efektif dari pada algoritme *Jaro-Winkler* karena pada beberapa percobaan *Rabin-Karp* cenderung mendapatkan persentase kemiripan yang lebih tinggi dengan rata-rata nilai persentase kesamaan 51% dan 35% untuk algoritme *Jaro-Winkler*. Dalam hal waktu pemrosesan dan untuk beberapa dokumen yang memiliki ukuran besar, *Rabin-Karp* jauh lebih cepat dari *Jaro-Winkler* dengan selang waktu untuk semua skenario adalah sekitar 0,389 menit.

Penelitian selanjutnya mengenai algoritme *Rabin-Karp* yang dilakukan oleh (Shivaji & S, 2015) dimana hasil yang didapat dari penelitian ini yaitu dengan nilai presisi mencapai 85% dan juga dapat meminimalisir kegagalan deteksi sekitar 10%. Selanjutnya penelitian yang dilakukan oleh (Rani & Kaur, 2018) menerapkan algoritme *Rabin-Karp* dengan *Binary Search Tree* (BST) untuk deteksi kesamaan dalam dokumen sansekerta. Penelitian ini menguji keakuratan dari kedua algoritme tersebut dengan ukuran *file* yang berbeda. Hasil dari penelitian ini menunjukkan bahwa *Rabin-Karp* dan BST bekerja cukup efisien untuk mendeteksi kesamaan karena dapat melakukan perhitungan dengan ukuran *file* yang berbeda dari 10 kb hingga 1000 kb, tetapi dari segi waktu untuk *file* ukuran besar dengan kalkulasi waktu tertinggi sebesar 9546 dalam *milliseconds* pada data yang berukuran 1000 kb.

Penelitian selanjutnya mengenai metode *Cosine similarity* yang dilakukan oleh (K.Popat, Desmukh, & Metre, 2017) tentang *clustering* dokumen berdasarkan ukuran dari *Cosine similarity*, hasil dari penelitian ini menunjukkan nilai entropi

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

yang terbaik ditujukan untuk *Cosine similarity* yaitu 0,4, 0,2, 0,3, 0,6, 0,3, hal ini menunjukkan *Cosine similarity* bekerja dengan baik dalam pengklasteran teks dibandingkan dengan metode *similarity* lainnya.

(Saptono, 2018) pada penelitiannya mengkombinasikan metode *Cosine similarity* dengan *Conditional Probability* untuk pendeteksian plagiarisme pada dokumen tesis. Penelitian ini menguji dengan 4 skenario dan nilai *threshold* yang digunakan adalah nilai dalam skenario 3, dengan nilai *threshold* 0,3 untuk *Conditional Probability* dan 0,2 untuk *Cosine similarity* dengan 54,28% untuk rata-rata presisi dan 100% untuk rata-rata *recall*. Hal ini menunjukkan bahwa rata-rata presisi dan *recall* lebih tinggi dibandingkan skenario 1, 2 dan 4. Maka penelitian ini menyimpulkan bahwa *Cosine similarity* dan *Conditional Probability* dapat diterapkan untuk pendeteksian kesamaan dokumen.

Penelitian terkait selanjutnya mengenai perbandingan metode *Cosine similarity*, *Jaccard Similarity*, dan *Euclidean Distance* yang diteliti oleh (Nurdiana, Jumadi, & Nursantika, 2016) dimana penelitian ini mendapatkan hasil bahwa metode *Cosine similarity* memiliki kemiripan yang tertinggi sebesar 41% dari metode *Jaccard* yang hanya 19% dan euclidean distance 40%. Metode *Cosine similarity* mempunyai konsep normalisasi panjang vektor data dengan membandingkan N-gram yang sejajar satu sama lain dari 2 pembanding. Sedangkan *Jaccard* hanya membandingkan isi N-gram dengan eksak dan *Euclidean Distance* tidak mempunyai konsep normalisasi panjang vektor data sehingga nilai akurasi dipengaruhi oleh panjang 2 data pembanding dengan harus menentukan nilai parameter *K*.

Berdasarkan dari beberapa penelitian yang telah dilakukan dan hasil dari kesimpulan yang telah diperoleh, maka pada penelitian ini akan menerapkan Algoritme *Rabin-Karp* dan *Cosine similarity* pada kasus pemeriksaan kesamaan dokumen tugas makalah mahasiswa dengan menampilkan *highlight* kalimat yang sama antar dokumen. Pemeriksaan kesamaan dokumen tugas makalah mahasiswa pada penelitian ini akan diuji berdasarkan mata kuliah Tata Tulis Karya Ilmiah (TTKI), dimana mata kuliah TTKI ini memiliki tugas dengan *output* dokumen makalah oleh setiap mahasiswa.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disebutkan diatas maka diadapatkan rumusan masalah yaitu:

1. Bagaimana menerapkan algoritme *Rabin-Karp* dan *Cosine similarity* untuk pemeriksaan kesamaan dokumen tugas makalah mahasiswa pada mata kuliah TTKI?
2. Bagaimana mengukur tingkat persentase kesamaan dalam penerapan algoritme *Rabin-Karp* dan *Cosine similarity* untuk dokumen tugas makalah mahasiswa pada mata kuliah TTKI?

1.3 Batasan Masalah

Batasan masalah dari penelitian ini adalah:

1. Dokumen yang diuji berupa dokumen teks digital berbahasa Indonesia dengan format *.doc* atau *.docx*.
2. Pemeriksaan Dokumen dilakukan untuk mendapatkan persentase *similarity* dari dokumen-dokumen yang diuji tersebut, tidak untuk mendapatkan *output* berupa keterangan plagiat atau tidak plagiat.
3. Sistem Pemeriksaan Kesamaan Dokumen ini tidak menggunakan *database* dalam pencarian data yang akan diuji.
4. Pemeriksaan kesamaan dokumen tugas makalah mahasiswa diuji berdasarkan tugas mata kuliah Tata Tulis Karya Ilmiah (TTKI) Teknik Informatika UIN Suska Riau.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Menerapkan algoritme *Rabin-Karp* untuk mendapatkan fitur dari dokumen tugas makalah mahasiswa pada mata kuliah TTKI.
2. Menerapkan dan menghitung tingkat persentase *Cosine similarity* untuk memeriksa kesamaan dokumen tugas makalah mahasiswa pada mata kuliah TTKI.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1.5 Sistematika Penulisan

Sistematika penulisan laporan Tugas Akhir ini terdiri dari pokok-pokok permasalahan yang dibahas pada masing-masing yang diuraikan menjadi beberapa bagian:

BAB I

PENDAHULUAN

Pada bab ini berisi uraian tentang latar belakang yang mendasari penelitian, rumusan masalah, tujuan penelitian, batasan masalah, serta sistematika penulisan.

BAB II

LANDASAN TEORI

Bab ini merupakan penjelasan mengenai studi pustaka terhadap teori-teori yang berhubungan dengan penelitian ini, analisa dan penulisan tugas akhir yang terdiri penjelasan tentang kesamaan dokumen, konsep *Text Mining*, proses *text preprocessing*, penerapan algoritme *Rabin-Karp*, kemudian penerapan metode *Cosine similarity*.

BAB III

METODOLOGI PENELITIAN

Pada bab ini akan dijelaskan bagaimana cara membuat kerangka penelitian serta metodologi penelitian yang digunakan untuk tugas akhir ini sampai selesai.

BAB IV

ANALISA DAN PERANCANGAN

Pada bab ini berisi tentang analisa yang dilakukan untuk membuat sistem yang sesuai dengan penelitian serta bagaimana bentuk perancangan sistem yang akan dibuat seperti, analisa kebutuhan data, analisa tahapan yang ada pada *text preprocessing*, analisa proses kerja algoritme *Rabin-Karp*, analisa metode *Cosine similarity* serta merancang aplikasi yang sesuai dengan kebutuhan penelitian.

BAB V

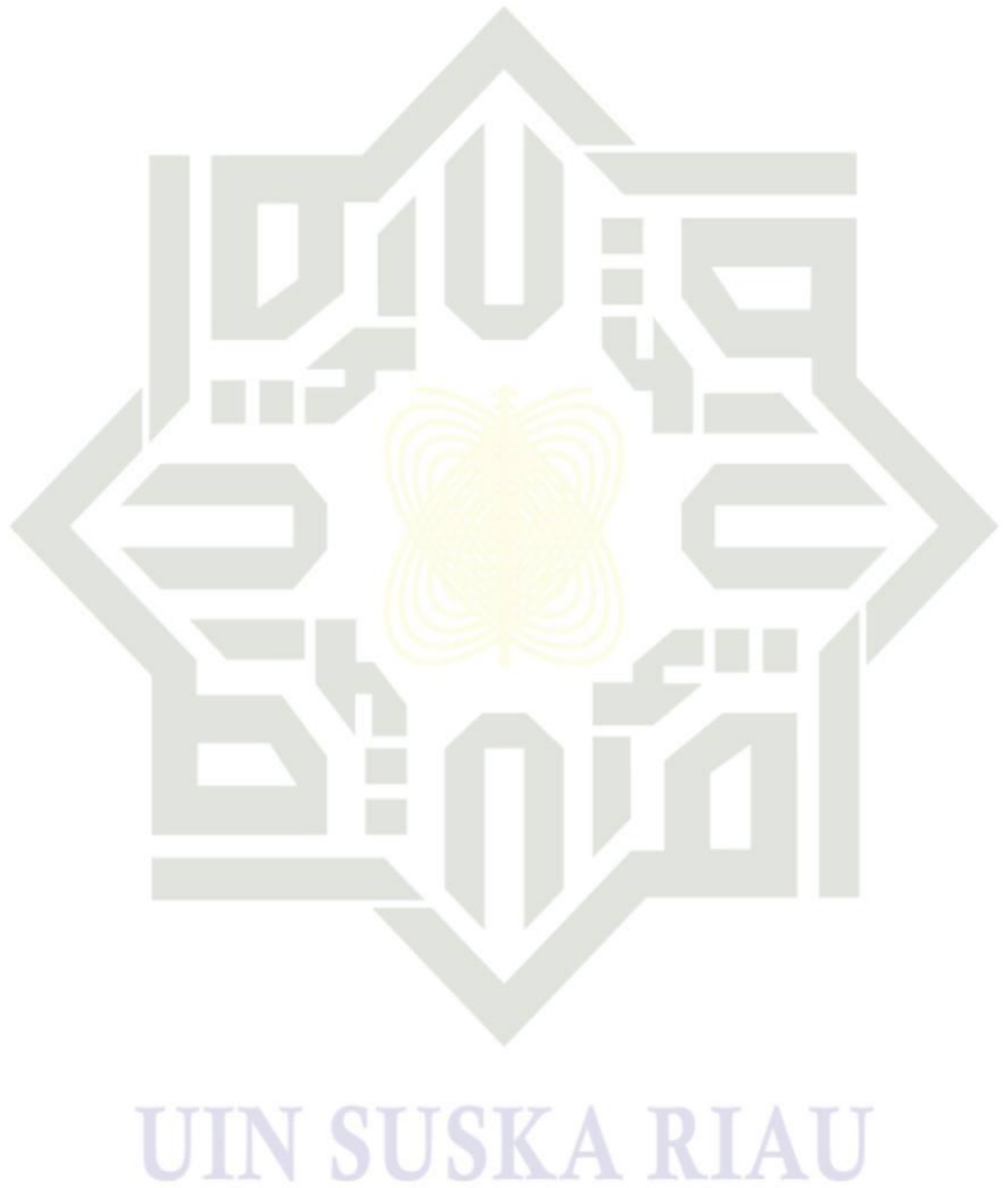
IMPLEMENTASI DAN PENGUJIAN

Pada bab ini akan dilakukan implementasi pada sistem yang dibangun sesuai dengan analisa dan perancangan yang telah dilakukan, serta melakukan pengujian apakah sistem berjalan sesuai fungsi atau tidak.

BAB VI

PENUTUP

Bab ini merupakan bab terakhir dari penelitian yang berisi kesimpulan dari penelitian yang telah dilakukan serta berisikan saran yang dapat membuat penelitian dapat dikembangkan.



- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

LANDASAN TEORI

2.1 Kesamaan Dokumen

Kesamaan dokumen (*document similarity*) merupakan fondasi dari sistem kecerdasan dalam pemrosesan data seperti *information retrieval*, klasifikasi teks, *clustering*, dan sebagainya. Kesamaan dokumen dapat digunakan sebagai alat pencarian informasi lain yang sejenis, sehingga dapat mempersingkat waktu. Kesamaan dokumen digunakan dengan melakukan perbandingan isi dari dokumen sehingga dari perbandingan tersebut dapat menghasilkan bobot atau nilai kemiripan dari dokumen yang dibandingkan, hal ini dapat membantu *user* dalam melakukan pekerjaan seperti pengelompokan dokumen atau juga dapat mendeteksi adanya tindakan plagiarisme dalam dokumen tersebut. Pendekatan kesamaan dokumen dikategorikan menjadi tiga kategori yaitu perbandingan teks lengkap, *keyword similarity* dan *fingerprint* (Stein & Eissen, 2005).

2.2 Metode Pengukuran Kesamaan Dokumen

Pendekatan pengukuran kesamaan terhadap dokumen teks terdapat 3 metode yaitu (Stein & Eissen, 2005):

1. Perbandingan Teks Lengkap

Penerapan dengan menggunakan metode ini dilakukan dengan cara melakukan perbandingan dengan semua isi dari dokumen. dokumen-dokumen yang akan dibandingkan adalah dokumen yang ada pada penyimpanan lokal. pada metode ini algoritme yang dapat diterapkan yaitu algoritme *Boyer Moore*, *Edit Distance*, *Lavenshtein Distance* dan *Brute Force*.

2. *Keyword Similarity*

Ide dari metode ini adalah melakukan ekstraksi kata kunci dari sebuah dokumen dan membandingkannya dengan kata kunci yang ada pada dokumen lain

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. *Fingerprint Analysis*

Pendekatan dengan menggunakan metode ini merupakan pendekatan yang digunakan untuk pendeteksian keakuratan kesamaan antar dokumen. metode ini memiliki prinsip kerja dengan menggunakan teknik *hashing*, dimana teknik *hashing* merupakan sebuah fungsi yang digunakan untuk melakukan konversi terhadap *string* menjadi bilangan. Algoritme yang digunakan pada metode ini antara lain seperti algoritme *Manber*, *Rabin-Karp*, dan *Winnowing*.

2.3 *Text Mining*

Text mining adalah proses pengetahuan intensif di mana seorang pengguna berinteraksi dengan koleksi dokumen dari waktu ke waktu dengan menggunakan seperangkat alat analisis (Feldman & Sanger, 2007). Definisi lain dari *text mining* yaitu proses *data mining* berupa teks yang memiliki sumber data berupa dokumen di mana bertujuan untuk melakukan pencarian terhadap kata-kata yang dapat mewakili isi dari sebuah dokumen untuk dilakukan proses analisa keterkaitan antara dokumen (J.Money, 2006). *Text mining* merupakan bentuk perluasan dari *data mining* atau *knowledge-discovery in database* (KDD) yang berfungsi untuk mengenali pola-pola menarik yang ada pada basis data skala besar (Nugroho, 2011). Secara garis besar proses yang terdapat pada *text mining* mengadopsi yang terdapat pada *data mining*, bahkan terdapat juga beberapa teknik-teknik dalam *data mining* yang terdapat pada proses *text mining*.

2.4 *Text Preprocessing*

Proses *preprocessing* dilakukan agar data bersih dari *noise*/gangguan serta menjadikan data ke bentuk yang lebih terstruktur untuk mempermudah proses pengolahan data selanjutnya. Proses yang dilakukan pada *preprocessing* seperti menghapus karakter yang tidak relevan dalam dokumen teks seperti tanda baca dan spasi, proses *preprocessing* juga dilakukan untuk mengkonversi huruf kapital menjadi non-kapital (Hasan, Wicaksana, & Hansun, 2018). Pada algoritme *Rabin-Karp*, tahapan *preprocessing* terdapat dalam proses yang dinamakan *case folding* yaitu penghilangan tanda baca dan mengubah teks menjadi kata-kata tanpa huruf kapital, *filtering* yaitu pembuangan karakter yang tidak penting, dan juga pada

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

proses *tokenizing/parsing K-gram* dimana merupakan proses membagi teks kedalam gram-gram yang ditentukan nilai *K-gram*nya.

2.5 Rabin-Karp

Algoritme *Rabin-Karp* merupakan algoritme pencocokan *string* yang ditemukan oleh Michael O Rabin dan Richard M Karp pada tahun 1987. Algoritme ini membandingkan nilai *hash* antara input *string* (m) dan *substring* dalam teks (n). Apabila nilai *hash* antara keduanya sama, maka akan dilakukan perbandingan sekali lagi terhadap karakter-karakternya, tetapi jika nilai *hash* tidak sama, maka *substring* akan bergeser ke kanan. Pergeseran dilakukan sebanyak (n-m) kali. Perhitungan nilai *hash* yang efisien selama pergeseran akan mempengaruhi kinerja dari algoritme ini (Sinaga & Hansun, 2018). Pencarian nilai *hash* pada algoritme *Rabin-Karp* dapat dilihat pada persamaan berikut ini:

$$H_{(c1...ck)} = c1 * b^{(k-1)} + c2 * b^{(k-2)} + \dots + c_{(k-1)} * b^k + c_k \quad (2.1)$$

Keterangan:

- H* : Nilai *hash*
- c* : Nilai ASCII karakter (desimal)
- b* : Basis (bilangan prima)
- k* : Banyak karakter (indeks karakter)

Basis bilangan harus dipilih bilangan prima yang cukup besar untuk mengurangi kemungkinan adanya dua *hash value* yang sama.

Karakteristik dari algoritme *Rabin-Karp* yaitu (Joshi, Upadhyay, & M, 2017):

1. Menerapkan fungsi *Hashing*.
2. Kompleksitas waktu yang digunakan pada fase *preprocessing* O(m).
3. Pada fase pencarian kompleksitasnya yaitu O(mn).
4. Waktu yang dibutuhkan O(m+n) dimana m adalah panjang *string* input dan n adalah jumlah iterasi yang dilakukan untuk menemukan solusi.

Berikut tahapan-tahapan dari algoritme *Rabin-Karp* yang akan dipaparkan pada sub-bab berikut ini.

2.5.1 Case folding

Case folding digunakan untuk penyeragaman karakter yang terdapat pada data. *Case folidng* Merupakan proses mengubah seluruh bentuk huruf besar yang ada pada teks menjadi huruf kecil, penghilangan karakter atau simbol-simbol dan

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

tanda baca lainnya yang tidak memiliki makna tertentu pada dokumen (Putra, Sujaini, & Pratiwi, 2015).

Contoh:

“UIN Suska Riau”

Kalimat tersebut akan diproses menjadi:

“uin suska riau”

2.5.2 Filtering

Filtering merupakan tahapan yang dilakukan setelah *case folding* yaitu mengambil kata-kata atau karakter-karakter penting pada sebuah dokumen.

Filtering dapat menggunakan algoritme *stoplist* yaitu membuang kata/karakter yang kurang penting) atau *wordlist* (menyimpan kata penting (Putra, Sujaini, & Pratiwi, 2015).

Pada penelitian ini akan menggunakan *stoplist* untuk membuang karakter-karakter yang dianggap tidak penting pada dokumen yang akan diuji seperti tanda titik (.), koma (,), garis miring (/) dan lain sebagainya.

Contoh:

“uin suska riau – teknik informatika.”

Kalimat tersebut akan diproses menjadi:

“uin suska riau teknik informatika”

2.5.3 Menentukan *K-gram*

Penentuan *K-gram* merupakan pembentukan pola kata dengan memecah kata menjadi potongan-potongan dimana setiap potongan mengandung karakter sebanyak *k*. *K-gram* merupakan sebuah metode yang diaplikasikan untuk pembangkitan kata atau karakter. Metode *K-gram* ini digunakan untuk mengambil potongan-potongan karakter huruf atau kata sejumlah *k* dari sebuah kata/kalimat yang secara kontinuitas dibaca dari teks sumber hingga akhir dari dokumen. (Nugroho, 2011). Berikut merupakan contoh *K-grams* dengan *k*=7.

- a. Diberikan kalimat “teknik informatika UIN suska riau”
- b. Lakukan penyeragaman karakter menjadi “teknik informatika uin suska riau”

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- c. Jika *K-gram* yang dilakukan per karakter huruf dengan rangkaian 7-grams yang diturunkan dari teks akan menjadi: “tekniki eknikin knikinf nikinfor ikinfor kinform informa nformat formati ormatik rmatika matikau atikau tikauin ikauins auinsus uinsusk insuska nsuskar suskari uskaria skariau”.
- d. Namun jika *K-gram* yang dilakukan perkata dengan rangkaian 3-grams yang diturunkan dari teks akan menjadi: “{teknik informatika uin} { informatika uin suska} { uin suska riau}.

Penggunaan nilai *K-gram* dapat disesuaikan dengan tingkat kesamaan dokumen yang ingin dicapai, semakin kecil nilai *K-gram* dapat mempertinggi tingkat akurasi dari kesamaan antar dokumen teks dan begitu juga sebaliknya.

2.5.4 Hashing

Setelah melakukan penentuan *K-gram* maka langkah selanjutnya adalah *hashing* terhadap seluruh pecahan *string* hasil dari proses parsing *K-gram*. *Hashing* itu sendiri adalah suatu cara untuk mentransformasi *string* menjadi suatu nilai yang unik (*hash value*) dengan panjang tertentu (*fixed-length*) yang berfungsi sebagai penanda *string* tersebut (Putra, Sujaini, & Pratiwi, 2015). Contoh sederhana dari *hashing* adalah:

Rabin,Michael;Karp,Richard

Menjadi:

1990 = Rabin,Michael;8822 = Karp,Richard

Pada contoh tersebut merupakan penerapan *hashing* dalam pencarian di *database*. Jika fungsi *hash* tidak dilakukan, maka pencarian akan dilakukan karakter per karakter pada sejumlah nama yang memiliki jumlah panjang yang berbeda-beda dan memiliki sebanyak 26 kemungkinan pada setiap karakter. Tetapi jika menggunakan fungsi *hash*, pencarian menjadi lebih efektif karena hanya akan membandingkan empat digit angka dengan 10 kemungkinan setiap angkanya.

Algoritme *Rabin-Karp* untuk pencarian pada pola yang panjang dan teks yang besar, algoritme ini menggunakan operasi *mod*, jika operasi *mod q* diterapkan, maka nilainya akan menjadi lebih kecil dari *q*. Berikut rumus matematis dari algoritme *Rabin-Karp* pada persamaan 2.2 berikut (Sukmana, Kusrini, & Sunyoto, 2018):

$$Hash(w[0 .. m-1]) = (w[0] * b^{m-1} + w[1] * b^{m-2} + \dots +$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

$$w[m-1] * b^0) \bmod q$$

$$Hash_{(K-gram\ 1)} = \sum hash(w_1) + \sum hash(w_n) \quad (2.2)$$

keterangan:

- b : Nilai bilangan basis
 w : Nilai ASCII karakter
 $h(w)$: Nilai *hash*
 p : Nilai bilangan prima (100007)
 m : Banyaknya karakter yang di *hashing*

Penggunaan *modulo* berbasis bilangan prima besar pada fungsi *hash* dikarenakan untuk mengurangi kemungkinan dua buah *corresponding number value* yang sama (Sinaga & Akbar, 2018).

Untuk mengetahui nilai ASCII dari suatu karakter, dapat dilihat pada tabel dibawah ini:

Tabel 2.1 Tabel ASCII

Karakter	Nilai ascii	Karakter	Nilai ascii	Karakter	Nilai ascii
A	65	a	97	0	48
B	66	b	98	1	49
C	67	c	99	2	50
D	68	d	100	3	51
E	69	e	101	4	52
F	70	f	102	5	53
G	71	g	103	6	54
H	72	h	104	7	55
I	73	i	105	8	56
J	74	j	106	9	57
K	75	k	107		
L	76	l	108		
M	77	m	109		
N	78	n	110		
O	79	o	111		
P	80	p	112		
Q	81	q	113		
R	82	r	114		
S	83	s	115		
T	84	t	116		
U	85	u	117		
V	86	v	118		
W	87	w	119		

Hak Cipta Dilindungi Undang-Undang

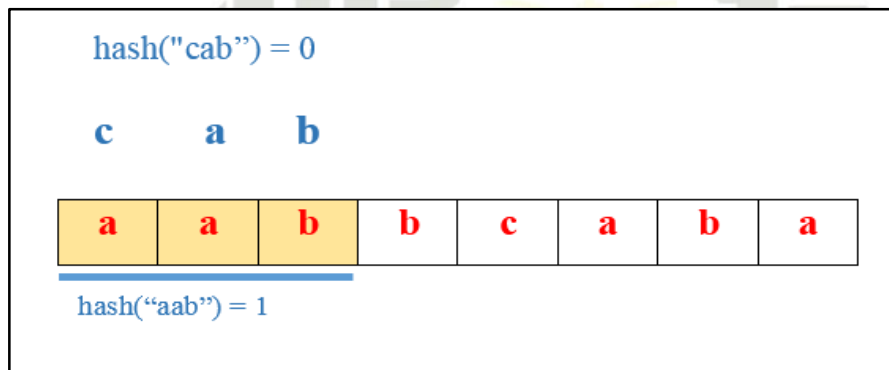
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

X	88	x	120		
Y	89	y	121		
Z	90	z	122		

2.5.5 Membandingkan Nilai Hash yang Sama

Berikut ini merupakan gambaran proses dari algoritme *Rabin-Karp* yang dirujuk dari (Sinaga & Hansun, 2018) untuk dilakukan perbandingan pada nilai *hash* yang sama terhadap *substring*.

Diberi “cab” sebagai input dan “aabbcab” sebagai teks. Fungsi *hash* yang digunakan misalnya akan menambahkan nilai keterututan setiap huruf dalam alfabet (a = 1, b = 2, dst) lalu modulo dengan 3. Nilai *hash* dari “cab” akan menjadi 0 dan tiga karakter pertama dalam teks yang merupakan “aab” akan menjadi 1 seperti yang ditunjukkan pada gambar 2.1.



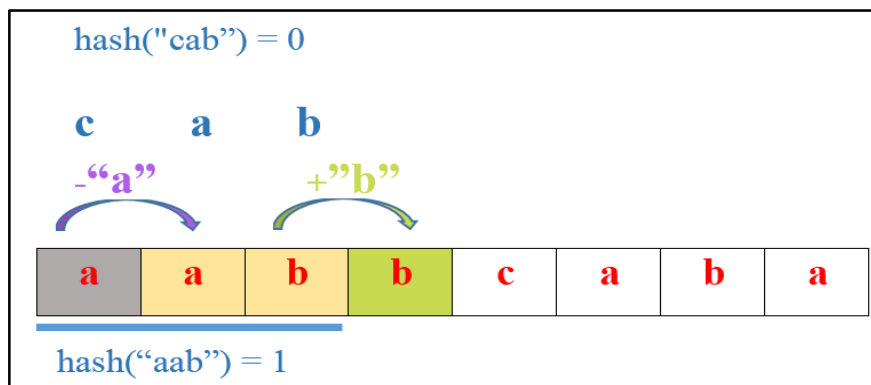
Gambar 2.1 Fingerprint awal

Gambar 2.1 menunjukkan perbandingan yang tidak sama, dan kemudian *substring* pada teks akan menggeser satu karakter ke kanan. Algoritme tidak akan menghitung ulang nilai *hash substring*, inilah yang disebut dengan *rolling hash*, yang mengurangi nilai pada karakter yang keluar lalu menambahkan dengan nilai karakter yang masuk, sehingga akan ada kompleksitas waktu pada setiap pergeseran (*shift*) yang dilakukan akan relatif lebih konstan yang relatif konstan (Sinaga & Hansun, 2018).

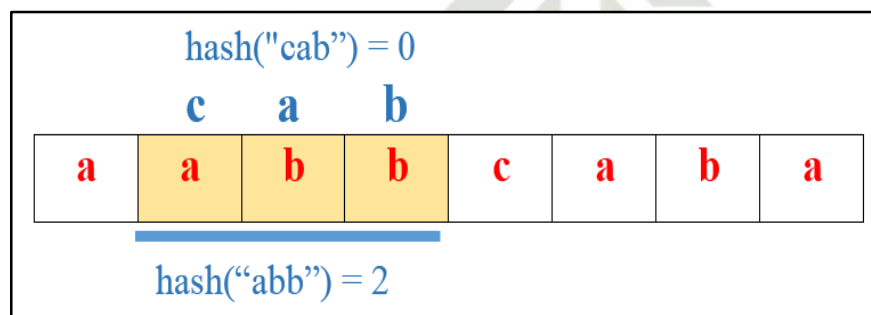
Setelah bergeser, nilai *hash* dari *fingerprint* “abb” ($\text{aab} = \text{aab} - \text{a} + \text{b}$) menjadi dua ($2 = 1 - 1 + 2$) seperti pada gambar 2.2 dan 2.3

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



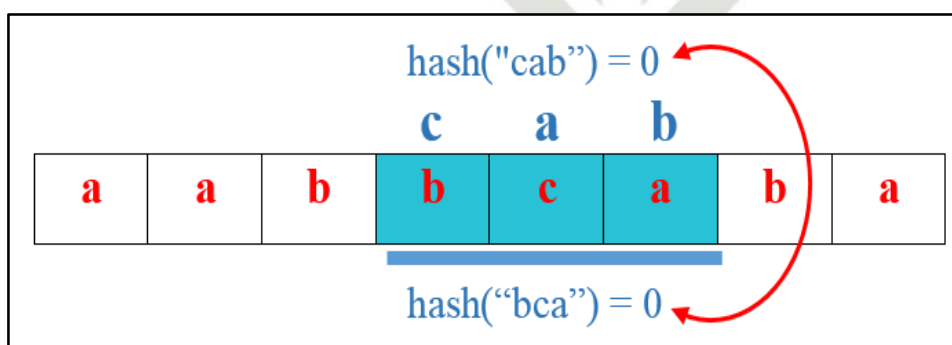
Gambar 2.2 Pergeseran *Fingerprint*



Gambar 2.3 Perbandingan Kedua

Gambar 2.3 menunjukkan hasil perbandingan juga tidak sama, sehingga harus dilakukan.

Pada perbandingan keempat, nilai *hash* yang diperoleh sama, karena nilai *hash* sama, maka dilakukan perbandingan *string* karakter per karakter, kemudian *string* dari karakter per karakter antara "bca" dan "cab" menunjukkan hasil bahwa kedua *string* tidak sama. Maka, *substring* bergeser ke kanan lagi, dapat dilihat pada gambar 2.4



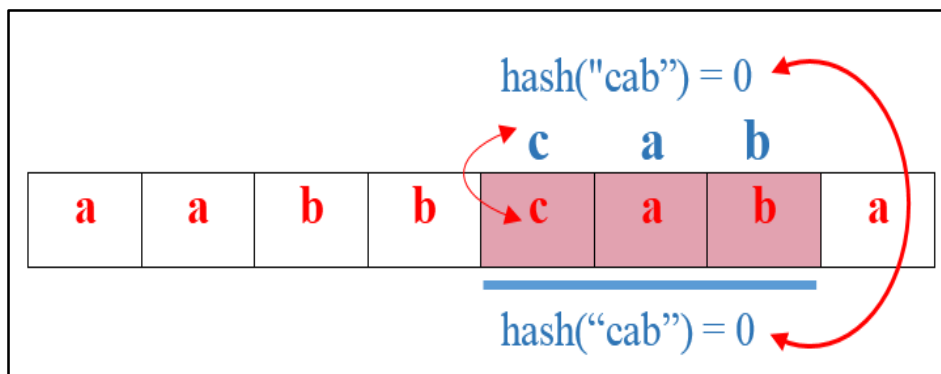
Gambar 2.4 perbandingan keempat (nilai *hash* yang sama)

Pada perbandingan kelima, nilai *hash* dan *string* dari karakter yang terbentuk sesuai, yang berarti solusi telah ditemukan. Dari perhitungan, diperlukan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

kompleksitas waktu yaitu $O(m + n)$ dimana m adalah panjang dari *string input* dan n adalah jumlah iterasi yang dilakukan untuk mencari solusi. Hasil ini dapat dilihat pada gambar 2.5 sebagai berikut.



Gambar 2.5 perbandingan kelima (*string* ditemukan)

2.6 Similarity

Pada penelitian ini, perhitungan similaritas dilakukan dengan metode *Cosine similarity* yang merupakan sebuah metode yang digunakan untuk perhitungan tingkat similaritas (kesamaan) antar dua buah objek. Perhitungan dengan metode ini berdasarkan pada *vector space similarity measure*, untuk menghitung similaritas pada dua buah objek dapat dimisalkan dengan D1 dan D2 yang dapat dinyatakan dalam dua buah *vector* menggunakan kata kunci dari dokumen sebagai ukuran.

Pada proses pencarian kesamaan dokumen yang bukan berdasarkan kata kunci, maka kata kunci diganti dengan kumpulan kata pada suatu dokumen yang akan dicari dokumen lain yang mirip dengan dokumen tersebut. *Query* dihasilkan dari kata-kata yang muncul di dokumen yang terpilih. Dibandingkan dengan metode perhitungan similaritas lainnya seperti *jaccard similarity*, metode *cosine* dinilai lebih efektif karena *Cosine similarity* memiliki konsep normalisasi pada panjang *vector* data dan membandingkan *N-gram* dari 2 pembandingan yang sejajar (Nurdiana, Jumadi, & Nursantika, 2016).

Berikut merupakan fungsi pada *Cosine Similarity* (Mardeana & Hidayah, 2016) :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

$$\text{Sim}(D_i, D_j) = \frac{\sum_{k=1}^d D_{ik} D_{jk}}{\sqrt{\sum_{k=1}^d (D_{ik})^2 * \sum_{k=1}^d (D_{jk})^2}} \quad (2.3)$$

Keterangan :

D_{ik} = *Fingerprint Frequency* k pada dokumen i

D_{jk} = *Fingerprint Frequency* k pada dokumen j

$D_{ik} D_{jk}$ = Jumlah *hash* yang sama pada D_i dan D_j

d = Dokumen

k = *Fingerprint Frequency*

2.7 Penelitian Terkait

Berikut merupakan penelitian yang pernah dilakukan sebelumnya mengenai kasus yang akan teliti yaitu terhadap pemeriksaan kesamaan dokumen, penerapan algoritme *Rabin-Karp* serta penerapan metode *Cosine Similarity*.

Tabel 2.2 Penelitian Terkait

No	Peneliti/tahun	Metode	Hasil
1	Er.Mohammad Shabaz, Er.Neha Kumari / 2017	<i>Rabin-Karp</i>	Pada penelitian tersebut mendapatkan hasil kesimpulan bahwa algoritme <i>Rabin-Karp</i> bekerja lebih baik dalam kompleksitas waktu yang lebih singkat
2	Sonawane Kiran Shivaji, Prabhudeva S / 2015	<i>Rabin-Karp</i>	Hasil nilai presisi mencapai 85% ke atas, dan juga dapat meminimalisir kegagalan deteksi sekitar 10%
3	Omkar sunil joshi, bhargavi R.upadhyay, supriya M / 2017	<i>Rabin-Karp</i>	Penggunaan algoritme <i>Rabin-Karp</i> memiliki kompleksitas waktu yang terbaik sejauh ini dibandingkan dengan algoritme knuth morns pratt.
4	Monika Rani, Harpreet Kaur / 2018	Rabin karp, binary search tree	Hasil yang didapatkan cukup efisien untuk mendeteksi kesamaan dokumen dalam waktu yang kurang efisien untuk <i>file</i> ukuran kecil hingga besar, 10 kb – 100 kb.

Hak Cipta Dilindungi Undang-Undang

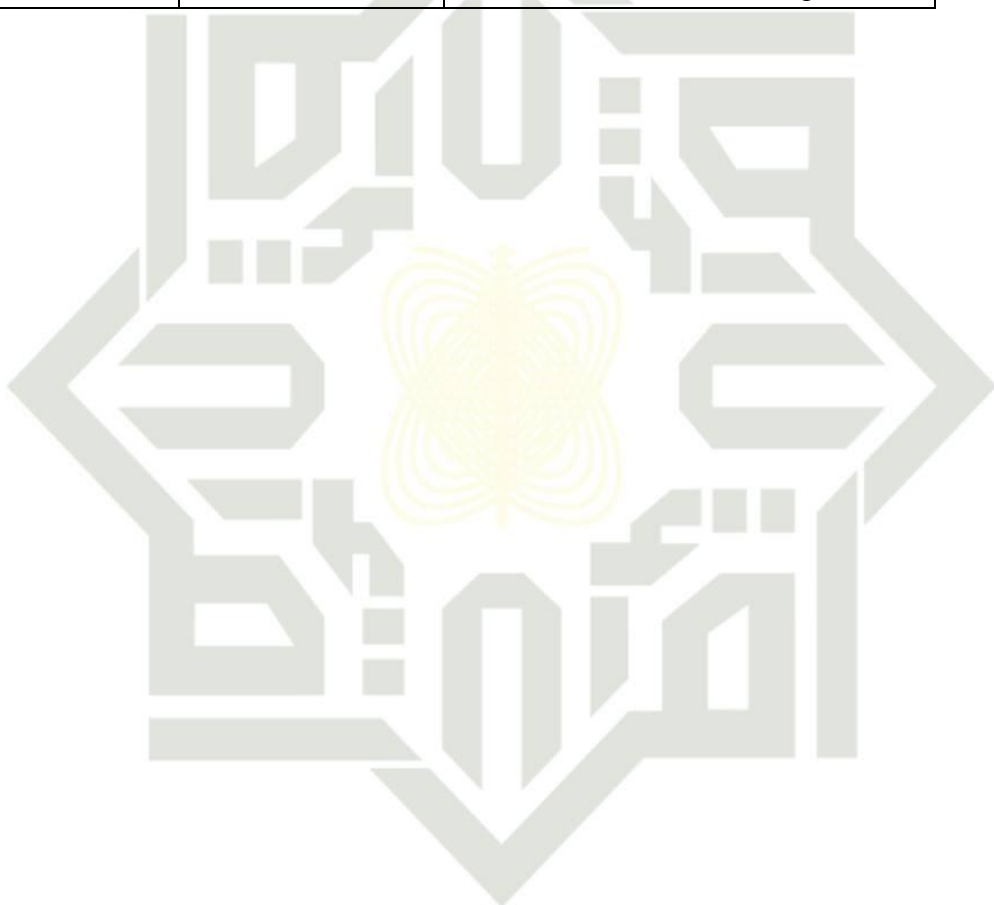
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti/tahun	Metode	Hasil
5	Deardo dibrianto sinaga dan seng hansun / 2018	<i>Rabin-Karp</i> , confix stripping algorithm	Hasil yang didapatkan dengan <i>Rabin-Karp</i> murni dengan kecepatan waktu pemrosesan 0.0123 detik dan tingkat <i>similarity</i> 89.1967%. proses <i>stemming</i> atau N-Gram juga dapat meningkatkan beberapa hasil tes dari waktu pemrosesan dan level <i>similarity</i>
6	Brinardi Leonardo, Seng Hansun / 2017	<i>Rabin-Karp</i> , <i>Jaro Winkler</i>	<i>Rabin-Karp</i> cenderung mendapatkan persentase kemiripan yang lebih tinggi dengan rata-rata nilai persentase kesamaan 51% dan 35% untuk algoritme <i>Jaro-Winkler</i> dan selang waktu pada <i>Rabin-Karp</i> dan <i>JaroWinkler</i> untuk semua skenario adalah sekitar 0.389 menit
7	Ristu Saptono, Heri Prasetyo, Ade Irawan / 2018	<i>Cosine similarity</i>	Penelitian ini mendapatkan hasil dari <i>threshold</i> 0,3 untuk <i>Conditional Probability</i> dan 0.2 untuk <i>Cosine similarity</i> dengan 54.28% untuk rata-rata presisi dan 100% untuk rata-rata recall
8	Ms.Shradda K.Popat, Ms.Vishaka A.Metre, Ms.Pramod B.Deshmukh / 2017	<i>Cosine similarity</i>	hasil dari penelitian ini menunjukkan nilai entropi yang terbaik ditujukan untuk <i>Cosine similarity</i> , hal ini menunjukkan <i>Cosine similarity</i> bekerja dengan baik dalam pengklasteran teks
9	Qamar mahmood, muhammad abdul qadir, muhammad tanvir afzal / 2013	Semantic social network analysis pada RDF citation graph	Hasil dari RDF citation graph dengan <i>similarity</i> 72.3345% dekat dengan <i>link node similarity</i>
10	Agung toto wibowo, kadek w.sudarmadi, Ari M.Barmawi / 2013	<i>Fingerprint</i> , Algoritme <i>winnowing</i>	Pada penelitian ini menunjukkan hasil dari <i>fingerprint</i> algoritme sebesar 92% dari <i>winnowing</i> sebesar 91.8%. pada tingkat relevansi dengann topik hasil menunjukkan algoritme <i>winnoiwn</i> g mendaptkan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Peneliti/tahun	Metode	Hasil
			kerelasi jangka yang lebih kuat 37.1% dibandingkan fingerprint 33.6%
11	Xiaojie yu, xiaojun chen, jinqiao shi/ 2017	LSA dan VSM	Hasil yang didapat dari penelitian ini yaitu LSA bekerja lebih baik dari pada VSM, dimana akurasi terbaik didapatkan dari KSA-150 dengan akurasi 0.86 dan VSM dengan 0.80



UIN SUSKA RIAU

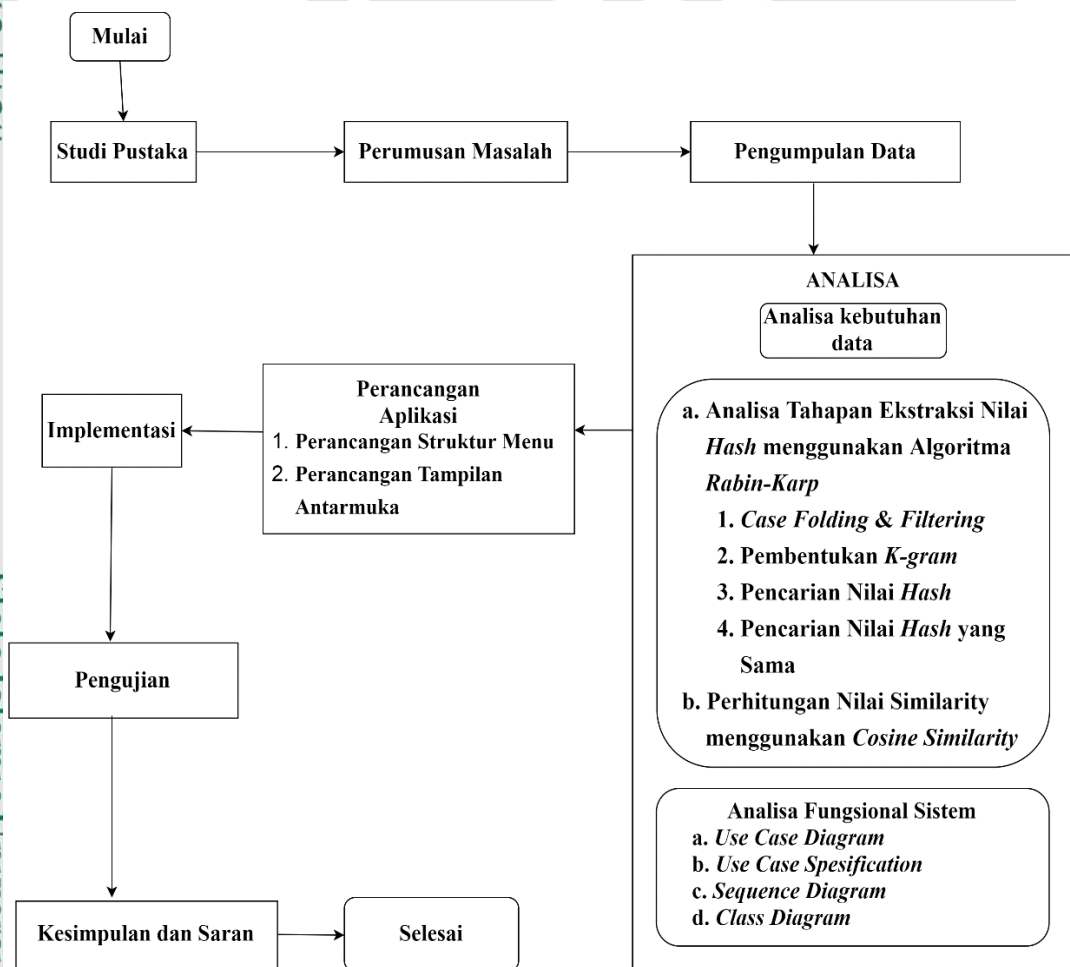
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB III

METODOLOGI PENELITIAN

Metodologi penelitian merupakan langkah-langkah atau tahapan yang dilakukan dalam penelitian. Metodologi penelitian ini bertujuan agar penelitian yang dilakukan sesuai dengan tujuan dari penelitian itu sendiri. Pada penelitian Tugas Akhir ini terdapat beberapa tahapan penelitian yang akan dilakukan seperti yang terlihat pada Gambar 3.1 dibawah ini.



Gambar 3.1 Tahapan Penelitian

3.1 Studi Pustaka

Studi pustaka merupakan tahapan paling awal dalam penelitian ini. Tahapan ini dilakukan untuk mencari referensi-referensi yang berkaitan dengan penelitian. Pencarian referensi-referensi tersebut dapat berupa dari jurnal-jurnal, buku-buku,

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

artikel-artikel dari media internet atau media lainnya yang berkaitan dengan *text mining* dan sistem pemeriksaan kesamaan dokumen, hal ini menjadi dasar dalam perumusan masalah yang akan diteliti.

3.2 Perumusan Masalah

Setelah dilakukan studi pustaka terkait pada penelitian, tahapan selanjutnya dilakukan perumusan masalah mengenai algoritme *Rabin-Karp* dengan menggunakan metode *Cosine Similarity* untuk pemeriksaan dokumen tugas makalah mahasiswa Teknik Informatika UIN Suska Riau.

3.3 Pengumpulan Data

Pada tahapan pengumpulan data dilakukan dengan melakukan wawancara ke beberapa dosen Teknik Informatika UIN Suska Riau serta mengumpulkan dokumen-dokumen tugas makalah mahasiswa khususnya pada mata kuliah TTKI. Dokumen berupa data teks berbahasa indonesia yang berbentuk digital dengan format *.doc* dan *.docx* untuk dilakukan pemeriksaan kesamaan dokumen antar tugas makalah mahasiswa..

3.4 Analisa

Setelah proses pengumpulan data selesai dilakukan, tahap berikutnya yang harus dilakukan adalah menganalisa hal-hal yang berkaitan dengan sistem yang akan dibangun. Tahapan analisa ini berhubungan dengan melakukan identifikasi kebutuhan dalam suatu penelitian. Beberapa tahapan analisa yang dilakukan yaitu analisa kebutuhan data, analisa proses kerja algoritme *Rabin-Karp* yang akan dihitung *similarity*-nya pada persamaan *Cosine Similarity*, serta menganalisa fungsional sistem.

3.4.1 Analisa Kebutuhan Data

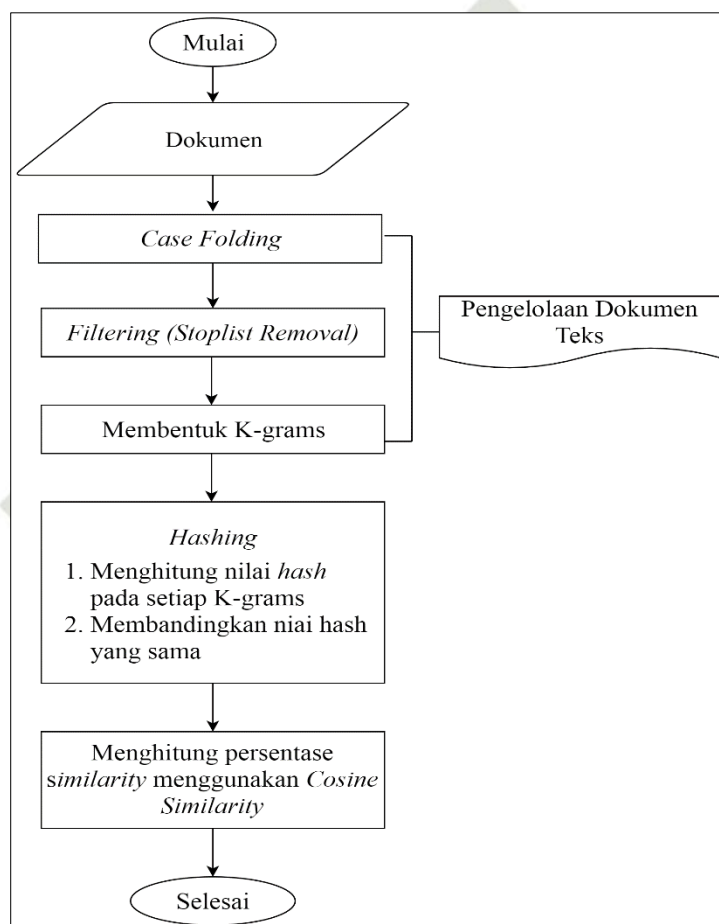
Tahapan ini merupakan tahapan yang dilakukan pada data-data yang sudah dikumpulkan yang akan diproses menggunakan algoritme *Rabin-Karp*, data-data yang digunakan adalah *softcopy* dokumen Tugas Makalah mahasiswa Teknik Informatika UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.4.2 Analisa Tahapan Ekstraksi Nilai *Hash* Menggunakan Algoritme *Rabin-Karp*

Tahapan ini dilakukan untuk mengidentifikasi langkah-langkah yang terdapat pada algoritme *Rabin-Karp* dan melakukan perhitungan kesamaan dengan *Cosine Similarity*. Langkah-langkah proses dalam algoritme *Rabin-Karp* hingga perhitungan dengan *Cosine Similarity* dapat dilihat pada *flowchart* berikut:



Gambar 3.2 Flowchart Proses *Rabin-Karp* dan *Cosine Similarity*

1. *Case folding*

Case folding digunakan untuk penyeragaman karakter yang terdapat pada data. Merupakan proses mengubah seluruh bentuk huruf besar yang ada pada teks menjadi huruf kecil yang terdapat pada dokumen.

2. *Filtering*

Filtering merupakan mengambil kata penting atau membuang kata/karakter-karakter yang tidak penting pada sebuah dokumen. Pada penelitian ini akan menggunakan *stoplist* untuk membuang karakter-

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

karakter yang dianggap tidak penting pada dokumen yang akan diuji seperti tanda titik (.), koma (,), garing miring (/) dan lain sebagainya.

3. Membentuk *K-gram*

Penentuan *K-gram* merupakan pembentukan pola kata dengan memecah kata menjadi potongan-potongan dimana setiap potongan mengandung karakter sebanyak *k*. *K-gram* merupakan sebuah metode yang diaplikasikan untuk pembangkitan kata atau karakter. Penggunaan *K-gram* dapat dapat disesuaikan dengan tingkat kesamaan yang ingin dicapai. Dengan menurunkan nilai *K-gram* dapat mempertinggi atau menambah akurasi hasil kesamaan yang diperoleh.

4. *Hashing*

Hashing adalah suatu cara untuk mentransformasi *string* menjadi suatu nilai yang unik (*hash value*) dengan panjang tertentu (*fixed-length*) yang berfungsi sebagai penanda *string* tersebut. Setelah melakukan proses tokenisasi dengan menentukan *K-gram*, maka setiap rangkaian *K-gram* akan dihitung nilai *hash* sebagai bobot dari fitur (*string*). Pada penelitian ini, akan dilakukan *hashing* pada beberapa dokumen untuk diuji. Setelah nilai *hash* dokumen-dokumen yang diuji berhasil diperoleh, maka akan dibandingkan nilai *hash* yang sama antara dokumen 1 dan dokumen 2.

5. Menghitung Kesamaan Menggunakan *Cosine Similarity*

Setelah proses pada algoritme *Rabin-Karp* selesai, maka setiap nilai *hash* yang sama antar dokumen akan dihitung tingkat similaritasnya. Perhitungan similaritas dilakukan dengan metode *Cosine Similarity* yang merupakan sebuah metode yang digunakan untuk perhitungan tingkat similaritas (kesamaan) antar dua buah objek dengan cara membandingkan setiap fitur yang telah diberi bobot menggunakan algoritme *Rabin-Karp*. Perhitungan dengan metode ini berdasarkan pada *vector space similarity measure*, untuk menghitung similaritas pada dua buah objek.

3.4.3 Analisa Fungsional Sistem

Analisa fungsional sistem merupakan analisa fungsi-fungsi yang dapat mendukung proses dari sistem. Pada penelitian ini, analisa fungsional sistem

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

menggunakan metode berorientasikan objek menggunakan UML (*Unified Modelling Language*).

1. *Use case Diagram*

Use case diagram menggambarkan segi fungsionalitas yang diberikan sebuah sistem kepada *user*, selain itu menunjukkan interaksi antara aktor-aktor dengan sistem yang akan dibangun.

2. *Use case Spesification*

Pada *use case spesification* ini setiap *use case* yang terdapat pada *use case diagram* harus dijelaskan alur kerjanya dengan menggunakan skenario.

3. *Class Diagram*

Class diagram menggambarkan struktur sistem dari segi pendefinisian kelas-kelas yang akan dibuat untuk membangun sistem. Kelas memiliki apa yang disebut atribut dan metode atau operasi.

4. *Sequence Diagram*

Sequence diagram menggambarkan kelakuan objek pada *use case* dengan mendeskripsikan waktu hidup objek atau *message* yang dikirimkan dan diterima antar objek. *Sequence diagram* yang harus digambar adalah minimal sebanyak pendefinisian *use case* yang memiliki proses sendiri atau yang penting semua *use case* yang didefinisikan interaksi jalannya sudah dicakup pada *sequence diagram*.

3.5 Perancangan Aplikasi

Setelah dilakukan proses analisa terhadap sistem yang akan dibangun, tahapan selanjutnya yang dilakukan yaitu melakukan perancangan pada sistem yang bertujuan memberi gambaran kepada pengguna terhadap sistem yang akan dibangun sesuai dengan analisa yang telah dilakukan. Berikut tahapan-tahapan dalam perancangan aplikasi:

1. Perancangan Struktur Menu

Melakukan perancangan terhadap menu-menu apa saja yang memungkinkan dibutuhkan pada sistem yang akan dibangun berdasarkan tujuan dari sistem itu sendiri.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Perancangan Tampilan Antarmuka

Melakukan perancangan terhadap sistem yang dibangun agar dapat digunakan dan dimengerti oleh pengguna.

3.6 Implementasi

Tahapan implementasi merupakan tahapan dimana sistem yang sudah dianalisa dan dirancang siap untuk dioperasikan. Tahapan ini bertujuan untuk mengetahui apakah sistem yang telah dirancang dan dibangun sesuai dengan tujuan yang ingin dicapai. Dalam melakukan implementasi pada sistem pemeriksaan kesamaan dokumen ini perlu adanya perangkat pendukung yang berupa perangkat *hardware* dan *software* yaitu sebagai berikut:

1. Perangkat keras (*hardware*) yang digunakan

Processor : Intel(R) Core(TM) i5-3210M CPU @ 2.50 GHz 2.50 GHz
RAM : 4,00 GB
System Type : 64-bit Operating System
Harddisk : 465,76 GB

2. Perangkat Lunak (*Software*)

Sistem Operasi : Microsoft Windows 7 Ultimate
Tools : NetBeans IDE 8.2
Browser : Google Chrome
Bahasa Pemrograman : JAVA
Software Pendukung : JavaFX SceneBuilder-8.5.0

3.7 Pengujian

Tahapan ini merupakan tahapan yang dilakukan untuk menguji tingkat keberhasilan sistem yang telah dibangun apakah sesuai dengan tujuan yang diharapkan atau tidak. Proses pengujian dilakukan dengan beberapa tahap sebagai berikut:

1. Pengujian fungsional algoritme dilakukan untuk membandingkan hasil perhitungan manual dan hasil output yang dihasilkan sistem dari proses-proses yang terdapat pada algoritme Pengujian tata letak kalimat yaitu pengujian yang dilakukan dengan mengubah posisi kalimat yang terdapat pada dokumen teks.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Pengujian dokumen yang sama yaitu pengujian yang dilakukan pada sistem dengan menginputkan dokumen yang sama.
3. Pengujian dokumen yang tidak memiliki kesamaan yaitu pengujian yang dilakukan pada sistem dengan bentuk dokumen baik topik atau isi pada dokumen yang tidak memiliki kesamaan.
4. Pengujian parameter K merupakan proses pengujian dengan memilih nilai k yang beragam untuk pengukuran tingkat *similarity* pada dokumen-dokumen yang diuji.
5. Pengujian banyak dokumen yaitu pengujian yang dilakukan dengan menginputkan *multiple file* dokumen pada sistem.
6. Pengujian perbandingan sistem dengan aplikasi lain. Perbandingan sistem dilakukan dengan aplikasi *Plagiarism Checker X* yang merupakan aplikasi pendeteksian plagiarisme pada dokumen.

3.8 Kesimpulan dan saran

Tahapan ini merupakan tahapan terakhir dari penelitian yang dilakukan, berisi tentang kesimpulan yang dapat diambil dari penelitian yang telah dilakukan dari tahapan analisa hingga sampai pada tahap pengujian. Setelah itu dari kesimpulan yang telah didapatkan dapat ditarik saran-saran yang diberikan untuk dilakukan pengembangan pada sistem yang dibangun agar menjadi lebih baik.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB IV

ANALISA DAN PERANCANGAN

Pada bab ini akan dilakukan analisa dan perancangan. Tahapan analisa terdiri dari analisa kebutuhan data, analisa, analisa proses ekstraksi nilai *hashing* menggunakan algoritme *Rabin Karp*, analisa tahapan perhitungan nilai kesamaan (*similarity*) menggunakan *Cosine Similarity* dan analisa fungsional sistem. Sedangkan pada tahapan perancangan terdapat beberapa perancangan yaitu perancangan struktur menu dan perancangan tampilan antarmuka.

4.1 Analisa

Analisa adalah suatu bentuk kegiatan yang bertujuan untuk mempelajari serta mengevaluasi suatu permasalahan yang dihadapi agar mencapai sebuah keputusan untuk mendapatkan solusi dari permasalahan tersebut. Pada penelitian ini akan dilakukan analisa terhadap kebutuhan data, analisa sistem pemeriksaan kesamaan dokumen, analisa algoritme *Rabin-Karp*, analisa proses ekstraksi nilai *hash*, dan analisa perhitungan persentase kesamaan.

4.1.1 Analisa Kebutuhan Data

Pada sistem pemeriksaan kesamaan isi dokumen ini data yang dibutuhkan untuk menjalankan fungsional dari sistem adalah data berupa dokumen-dokumen makalah mahasiswa Teknik Informatika Uin Suska Riau. Adapun makalah yang digunakan yaitu makalah mahasiswa dengan mata kuliah Tata Tulis Karya Ilmiah (TTKI). Pengujian sistem akan dilakukan pada dokumen-dokumen TTKI ini, dimana format dokumen yaitu berupa *.doc* dan *.docx*.

4.1.2 Analisa Proses Ekstraksi Nilai *Hashing* Menggunakan Algoritme *Rabin-Karp* Pada Dokumen 1

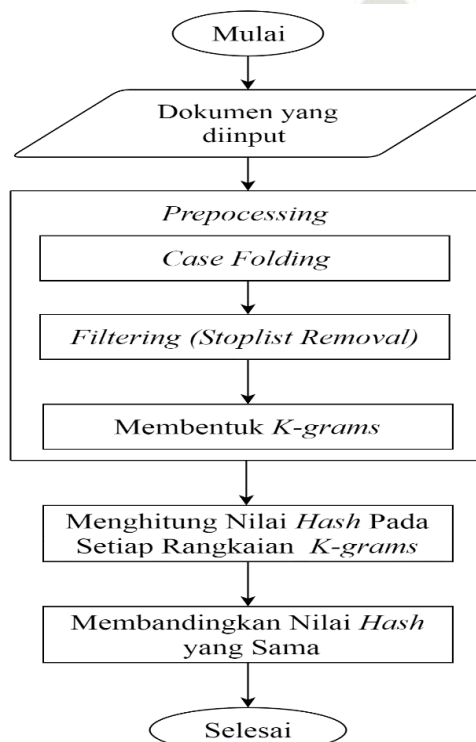
Pada tahapan ini, dilakukan analisa terhadap proses ekstraksi nilai *hash* menggunakan algoritme *Rabin-Karp*. Proses dilakukan dengan melakukan tahapan *preprocessing* seperti *case folding*, *filtering*, dan menentukan nilai *K-gram* serta basis bilangan primanya. Kemudian dari setiap hasil *K-gram* akan dicari nilai *hashing* dari setiap dokumen-dokumen yang diuji. Teknik *hashing* ini merupakan teknik yang digunakan untuk mengkonversi setiap *string* menjadi bilangan. Setiap

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

nilai *hashing* dari kedua data akan dilakukan perbandingan untuk mencari nilai *hash* yang sama sebagai *fingerprint*. Jika *hashing* yang sama ditemukan dari kedua dokumen tersebut, maka hasil *hashing* yang sama tersebut akan dihitung tingkat persentase *similarity* (kesamaan).

Lebih jelasnya, berikut ini merupakan tahapan-tahapan dari algoritme *Rabin-Karp*:



Gambar 4.1 Flowchart algoritme *Rabin-Karp*

Berdasarkan dari *flowchart* diatas, tahapan-tahapan algoritme *Rabin-Karp* yang dilakukan dalam perancangan sistem ini yaitu sebagai berikut:

1. Merubah huruf kapital menjadi huruf kecil.
2. Menghilangkan tanda baca.
3. Membagi teks kedalam bentuk *K-gram*, dimana nilai *k* merupakan nilai parameter yang dipilih oleh *user*.
4. Menghitung nilai *hash* dari setiap rangkaian *K-gram*.
5. Membandingkan nilai *hash* yang sama.

Pada tahapan ini, inputan berupa dokumen teks yang mempunyai ekstensi *.doc* dan *.docx*, dapat di inisialisasikan sebagai dokumen 1. Setelah itu, sistem akan

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

memproses dokumen tersebut sesuai dengan tahapan-tahapan yang terdapat pada algoritme *Rabin-Karp*.

Pertama kali proses yang dilakukan oleh sistem adalah membaca *file* teks yang diinputkan oleh *user*. Dari dokumen yang telah diinputkan oleh *user*, sistem akan melakukan pengecekan terhadap dokumen apakah tipe *file* tersebut berupa *doc* atau *.docx*. Berdasarkan dari algoritme *Rabin-Karp*, parameter yang harus diinput atau dipilih selanjutnya yaitu nilai *K* yang digunakan untuk proses tokenisasi.

Setelah sistem mendapatkan informasi dari dokumen yang telah diinputkan, sistem akan masuk ke tahap *preprocessing*. Pada tahap ini akan dilakukan beberapa proses, yaitu *case folding* (mengubah semua huruf kapital menjad huruf kecil, *filtering* (penghilangan tanda baca yang kurang penting), kemudian proses *tokenizing* (pembentukan rangkaian *K-gram*). Setelah itu dilakukan proses *hashing* untuk mengetahui nilai *hash* di setiap kalimat yang terdapat pada dokumen tersebut. Berikut merupakan contoh dokumen 1 yang akan diuji.

Tabel 4.1 Contoh Dokumen 1

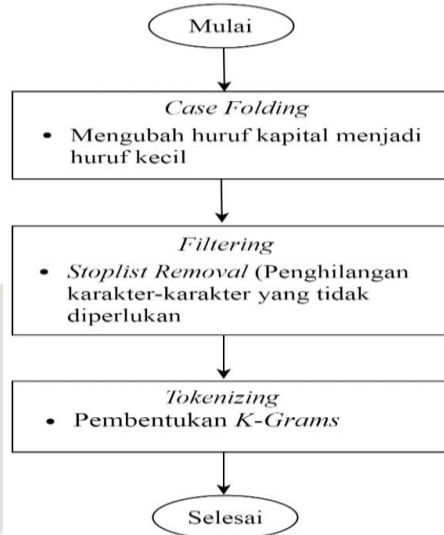
Motivasi berasal dari kata “motif” yang berarti daya upaya yang mendorong seseorang untuk melakukan sesuatu. Dorongan dasar yang menggerakkan tingkah laku manusia disebut sebagai motivasi. Motivasi adalah kekuatan yang muncul dari dalam diri seseorang untuk mendapatkan tujuan yang diinginkan dan berusaha melakukan perubahan tingkah laku yang lebih baik demi memenuhi keinginannya. Semakin besar dorongan dari dalam diri seseorang maka semakin besar pula usaha yang dilakukannya untuk mencapai tujuan.

Motivasi sangat diperlukan dalam proses belajar sebab seseorang tidak akan melakukan aktivitas belajar jika tidak mempunyai motivasi dalam belajar. Motivasi menimbulkan gairah, semangat dan merasa senang untuk belajar. Semakin besar motivasi seseorang maka semakin besar pula energi yang dimilikinya untuk belajar.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.1.2.1 Tahapan *Preprocessing*

Pada tahapan *preprocessing*, akan digambarkan dengan *flowchart* dibawah ini:

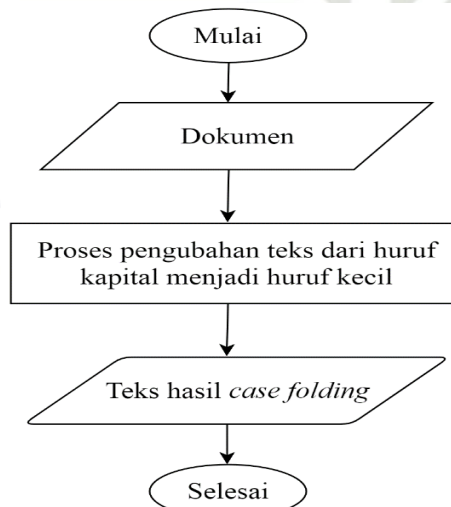


Gambar 4.2 Flowchart Tahapan *Preprocessing*

Tahapan *Preprocessing* harus dilalui agar proses perhitungan nilai *hash* dapat dilakukan pada kedua dokumen yang akan dilakukan pengujian, yaitu pada dokumen 1 dan dokumen 2. Terdapat beberapa proses yang akan dilakukan oleh sistem pada dokumen yang diinputkan di tahapan *preprocessing* ini yaitu diantaranya *case folding*, *filtering*, dan *tokenizing*.

4.1.2.1.1 Proses *Case folding*

Proses *case folding* akan digambarkan pada *flowchart* dibawah ini:



Gambar 4.3 Flowchart Proses *Case folding*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Case folding merupakan proses mengubah semua huruf kapital menjadi huruf kecil pada dokumen. Proses ini merupakan tahap pertama yang akan dilakukan dari serangkaian tahapan yang ada pada *preprocessing*. Contoh *case folding* dapat dilihat sebagai berikut.

Tabel 4.2 Dokumen 1 Setelah Melalui Proses *Case folding*

motivasi berasal dari kata “motif” yang berarti daya upaya yang mendorong seseorang untuk melakukan sesuatu. dorongan dasar yang menggerakkan tingkah laku manusia disebut sebagai motivasi. motivasi adalah kekuatan yang muncul dari dalam diri seseorang untuk mendapatkan tujuan yang diinginkan dan berusaha melakukan perubahan tingkah laku yang lebih baik demi memenuhi keinginannya. semakin besar dorongan dari dalam diri seseorang maka semakin besar pula usaha yang dilakukannya untuk mencapai tujuan.

motivasi sangat diperlukan dalam proses belajar sebab seseorang tidak akan melakukan aktivitas belajar jika tidak mempunyai motivasi dalam belajar. motivasi menimbulkan gairah, semangat dan merasa senang untuk belajar. semakin besar motivasi seseorang maka semakin besar pula energi yang dimilikinya untuk belajar.

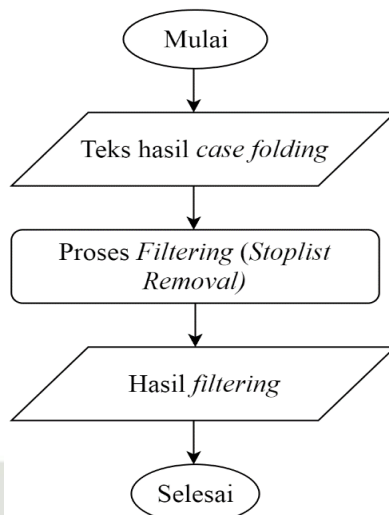
Berdasarkan contoh proses *case folding* tersebut, kalimat yang terdapat huruf kapital seperti “Motivasi”, “Dorongan”, “Semakin” akan dilakukan proses *case folding* dengan menjadikan kata-kata tersebut menjadi huruf kecil seperti kata “Motivasi” menjadi “motivasi”, kata “Dorongan” menjadi “dorongan”, dan kata “Semakin” menjadi “semakin”. Maka kata-kata yang terdapat huruf kapital pada sebuah dokumen akan menjadi huruf kecil secara keseluruhan.

4.1.2.1.2 Proses Filtering

Berikut merupakan *flowchart* dari *filtering*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4.4 Flowchart Proses Filtering

Setelah dilakukan proses *case folding*, maka tahapan selanjutnya masuk kepada proses *filtering*, dimana proses *filtering* pada sistem ini merupakan proses penghilangan tanda baca yang kurang penting dengan menggunakan algoritme *stop list* (membuang tanda baca yang kurang penting). Berikut contoh penerapan *filtering* dapat dilihat pada contoh dibawah ini.

Tabel 4.3 Dokumen 1 Setelah melalui Proses Filtering (Stoplist Removal)

motivasi berasal dari kata motif yang berarti daya upaya yang mendorong seseorang untuk melakukan sesuatu dorongan dasar yang menggerakkan tingkah laku manusia disebut sebagai motivasi motivasi adalah kekuatan yang muncul dari dalam diri seseorang untuk mendapatkan tujuan yang diinginkan dan berusaha melakukan perubahan tingkah laku yang lebih baik demi memenuhi keinginannya semakin besar dorongan dari dalam diri seseorang maka semakin besar pula usaha yang dilakukannya untuk mencapai tujuan motivasi sangat diperlukan dalam proses belajar sebab seseorang tidak akan melakukan aktivitas belajar jika tidak mempunyai motivasi dalam belajar motivasi menimbulkan gairah semangat dan merasa senang untuk belajar semakin besar motivasi seseorang maka semakin besar pula energi yang dimilikinya untuk belajar

Pada proses *filtering* ini, tanda baca seperti titik (.), koma (,), tanda seru (!), tanda tanya (?), titik dua dan titik dua koma (:)(;), garis miring (/), kutip dua (“) dan tanda baca lainnya akan dihilangkan. Contoh *filtering* yang terdapat pada

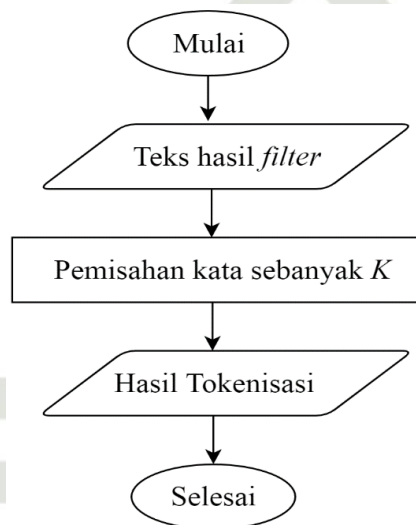
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dokumen 1 diatas terdapat beberapa tanda baca yang dianggap kurang penting seperti tanda *titik* (.), kutip dua (“), dan *koma* (,), maka melalui proses *filtering* tanda baca tersebut akan dihilangkan pada setiap dokumen, dan kalimat-kalimat tersebut akan menjadi satu kalimat yang utuh.

4.1.2.1.3 Proses pembentukan *K-gram*

Tahapan pembentukan *K-gram* akan digambarkan pada *flowchart* berikut ini:



Gambar 4.5 Flowchart Pembentukan *K-gram*

Tahapan *tokenizing* (pembentukan *K-gram*) merupakan tahapan ketiga dari proses *preprocessing* setelah *case folding* dan *filtering*. Pembentukan *K-gram* merupakan proses *tokenisasi* atau pemotongan kata guna untuk membentuk pola kata, dimana pola katanya dalam bentuk *gram* dengan panjang *k*.

Lebih jelasnya, pada tabel 4.4 berikut ini merupakan contoh proses pembentukan *K-gram* dengan melakukan pemotongan sebanyak *k*. Pada contoh ini proses pemotongan dilakukan dengan $k = 4$.

Tabel 4.4 Hasil *K-gram* dokumen 1

No.	Hasil <i>K-gram</i>	No.	Hasil <i>K-gram</i>
1	motivasi berasal dari kata	54	dorongan dari dalam diri
2	berasal dari kata motif	55	dari dalam diri seseorang
3	dari kata motif yang	56	dalam diri seseorang maka
4	kata motif yang berarti	57	diri seseorang maka semakin
5	motif yang berarti daya	58	seseorang maka semakin besar
6	yang berarti daya upaya	59	maka semakin besar pula
7	berarti daya upaya yang	60	semakin besar pula usaha

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No.	Hasil K-gram	No.	Hasil K-gram
8	daya upaya yang mendorong	61	besar pula usaha yang
9	upaya yang mendorong seseorang	62	pula usaha yang dilakukannya
10	yang mendorong seseorang untuk	63	usaha yang dilakukannya untuk
11	mendorong seseorang untuk melakukan	64	yang dilakukannya untuk mencapai
12	seseorang untuk melakukan sesuatu	65	dilakukannya untuk mencapai tujuan
13	untuk melakukan sesuatu dorongan	66	untuk mencapai tujuan motivasi
14	melakukan sesuatu dorongan dasar	67	mencapai tujuan motivasi sangat
15	sesuatu dorongan dasar yang	68	tujuan motivasi sangat diperlukan
16	dorongan dasar yang menggerakkan	69	motivasi sangat diperlukan dalam
17	dasar yang menggerakkan tingkah	70	sangat diperlukan dalam proses
18	yang menggerakkan tingkah laku	71	diperlukan dalam proses belajar
19	menggerakkan tingkah laku manusia	72	dalam proses belajar sebab
20	tingkah laku manusia disebut	73	proses belajar sebab seseorang
21	laku manusia disebut sebagai	74	belajar sebab seseorang tidak
22	manusia disebut sebagai motivasi	75	sebab seseorang tidak akan
23	disebut sebagai motivasi motivasi	76	seseorang tidak akan melakukan
24	sebagai motivasi motivasi adalah	77	tidak akan melakukan aktivitas
25	motivasi motivasi adalah kekuatan	78	akan melakukan aktivitas belajar
26	motivasi adalah kekuatan yang	79	melakukan aktivitas belajar jika
27	adalah kekuatan yang muncul	80	aktivitas belajar jika tidak
28	kekuatan yang muncul dari	81	belajar jika tidak mempunyai
29	yang muncul dari dalam	82	jika tidak mempunyai motivasi
30	muncul dari dalam diri	83	tidak mempunyai motivasi dalam
31	dari dalam diri seseorang	84	mempunyai motivasi dalam belajar
32	dalam diri seseorang untuk	85	motivasi dalam belajar motivasi
33	diri seseorang untuk mendapatkan	86	dalam belajar motivasi menimbulkan
34	seseorang untuk mendapatkan tujuan	87	belajar motivasi menimbulkan gairah

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

No.	Hasil <i>K-gram</i>	No.	Hasil <i>K-gram</i>
35	untuk mendapatkan tujuan yang	88	motivasi menimbulkan gairah semangat
36	mendapatkan tujuan yang diinginkan	89	menimbulkan gairah semangat dan
37	tujuan yang diinginkan dan	90	gairah semangat dan merasa
38	yang diinginkan dan berusaha	91	semangat dan merasa senang
39	diinginkan dan berusaha melakukan	92	dan merasa senang untuk
40	dan berusaha melakukan perubahan	93	merasa senang untuk belajar
41	berusaha melakukan perubahan tingkah	94	senang untuk belajar semakin
42	melakukan perubahan tingkah laku	95	untuk belajar semakin besar
43	perubahan tingkah laku yang	96	belajar semakin besar motivasi
44	tingkah laku yang lebih	97	semakin besar motivasi seseorang
45	laku yang lebih baik	98	besar motivasi seseorang maka
46	yang lebih baik demi	99	motivasi seseorang maka semakin
47	lebih baik demi memenuhi	100	seseorang maka semakin besar
48	baik demi memenuhi keinginannya	101	maka semakin besar pula
49	demi memenuhi keinginannya semakin	102	semakin besar pula energy
50	memenuhi keinginannya semakin besar	103	besar pula energi yang
51	keinginannya semakin besar dorongan	104	pula energi yang dimilikinya
52	semakin besar dorongan dari	105	energi yang dimilikinya untuk
53	besar dorongan dari dalam	106	yang dimilikinya untuk belajar

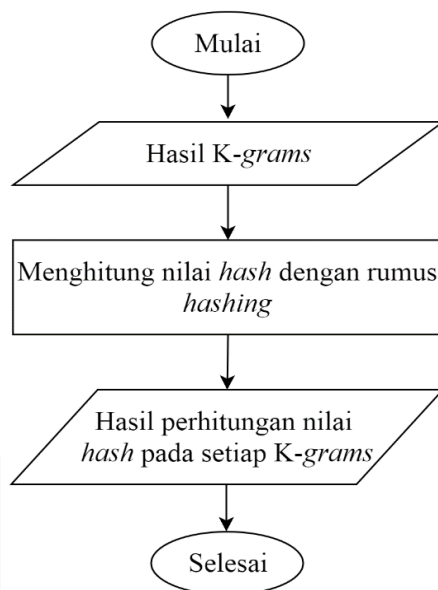
Pada contoh pembentukan *K-gram* diatas dilakukan proses pembentukan pola kata dalam bentuk *gram* dengan panjang kalimat $k = 4$. Maka setiap kalimat pada dokumen 1 berubah dalam bentuk rangkaian *K-gram* (token-token).

4.1.2.1.4 Proses Perhitungan Nilai Hash

Proses pencarian nilai *hashing* akan digambarkan pada *flowchart* dibawah ini:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4.6 Flowchart Proses Hashing

Setelah dilakukan tahapan *preprocessing* yaitu dengan *case folding*, *filtering*, dan membentuk pola kalimat menjadi rangkaian *gram-gram* sepanjang k karakter, maka tahapan selanjutnya yaitu melakukan proses pencarian nilai *hash* pada setiap rangkaian *K-gram* yang telah terbentuk.

Pada proses *hashing* ini berfungsi untuk mengkonvert setiap *string* menjadi bilangan. Dengan cara mengalikan nilai ASCII dengan huruf hasil *K-gram* dengan basis bilangan tertentu yang mana basis bilangan merupakan bilangan prima dan algoritme ini menggunakan operasi *mod* untuk pola yang panjang dan teks yang besar, dengan menggunakan persamaan yang ada pada (2.2). Pada algoritme *Rabin-Karp* mempunyai konsep jika dua buah *string* sama, maka *hash value*-nya pasti sama.

Berikut merupakan contoh perhitungan nilai *hash* pada dokumen 1 dari *K-gram* {motivasi berasal dari kata} dengan menggunakan $k = 4$, basis bilangan $b = 11$ dan modulo $q = 100007$

Perhitungan untuk Kalimat 1:

$$\text{Hash}(w[0 \dots m-1]) = (w[0] * b^{m-1} + w[1] * b^{m-2} + \dots + w[m-1] * b^0) \bmod q$$

$$\begin{aligned} H_{(\text{motivasi})} &= (\text{ascii}(m) * 11^{(7)} + \text{ascii}(o) * 11^{(6)} + \text{ascii}(t) * 11^{(5)} + \text{ascii}(i) * 11^{(4)} \\ &\quad + \text{ascii}(v) * 11^{(3)} + \text{ascii}(a) * 11^{(2)} + \text{ascii}(s) * 11^{(1)} + \text{ascii}(i) * 11^{(0)}) \\ &\quad \bmod 100007 \end{aligned}$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

$$\begin{aligned}
 &= (109 * 19.487.171 + 111 * 1.771.561 + 116 * 161.051 + 105 * 14.641 + 118 * 1.331 + 97 * 121 + 115 * 11 + 105 * 1) \bmod_{100007} \\
 &= (2.124.101.639 + 196.643.271 + 18.681.916 + 1.537.305 + 157.058 + 11.737 + 1.265 + 105) \bmod_{100007} \\
 &= 2.341.134.296 \bmod_{100007} \\
 &= 70.433 \\
 H_{(berasal)} &= (\text{ascii}(b) * 11^{(6)} + \text{ascii}(e) * 11^{(5)} + \text{ascii}(r) * 11^{(4)} + \text{ascii}(a) * 11^{(3)} + \text{ascii}(s) * 11^{(2)} + \text{ascii}(a) * 11^{(1)} + \text{ascii}(l) * 11^{(0)}) \bmod_{100007} \\
 &= (98 * 1.771.561 + 101 * 161.051 + 114 * 14.641 + 97 * 1.331 + 115 * 121 + 97 * 11 + 108 * 1) \bmod_{100007} \\
 &= (173.612.978 + 16.266.151 + 1.669.074 + 129.107 + 13.915 + 1.067 + 108) \bmod_{100007} \\
 &= 191.692.400 \bmod_{100007} \\
 &= 78.988 \\
 H_{(dari)} &= (\text{ascii}(d) * 11^{(3)} + \text{ascii}(a) * 11^{(2)} + \text{ascii}(r) * 11^{(1)} + \text{ascii}(i) * 11^{(0)}) \bmod_{100007} \\
 &= (100 * 1.331 + 97 * 121 + 114 * 11 + 105 * 1) \bmod_{100007} \\
 &= (133.100 + 11.737 + 1.254 + 105) \bmod_{100007} \\
 &= 146.196 \bmod_{100007} \\
 &= 46.189 \\
 H_{(kata)} &= (\text{ascii}(k) * 11^{(3)} + \text{ascii}(a) * 11^{(2)} + \text{ascii}(t) * 11^{(1)} + \text{ascii}(a) * 11^{(0)}) \bmod_{100007} \\
 &= (107 * 1.331 + 97 * 121 + 116 * 11 + 97 * 1) \bmod_{100007} \\
 &= (142.417 + 11.737 + 1.276 + 97) \bmod_{100007} \\
 &= 155.527 \bmod_{100007} \\
 &= 55.520
 \end{aligned}$$

$$\begin{aligned}
 \Sigma Hash_1 &= 70.433 + 78.988 + 46.189 + 55.520 \\
 &= 251.130
 \end{aligned}$$

Perhitungan nilai *hash* dari hasil *K-gram* selanjutnya dapat dilihat dari tabel 4.5 yang dirujuk pada lampiran B yang merupakan contoh perhitungan manual nilai *hash* pada token-token berikutnya..

Tabel 4.5 Nilai *hash* setiap *K-gram* dokumen 1

No.	Hasil <i>K-gram</i>	Nilai <i>Hash</i>	No.	Hasil <i>K-gram</i>	Nilai <i>Hash</i>
1	motivasi berasal dari kata	251.130	54	dorongan dari dalam diri	154.500
2	berasal dari kata motif	239.481	55	dari dalam diri seseorang	183.121
3	dari kata motif yang	234.587	56	dalam diri seseorang maka	195.015
4	kata motif yang berarti	267.471	57	diri seseorang maka semakin	207.011
5	motif yang berarti daya	258.209	58	seseorang maka semakin besar	244.094
6	yang berarti daya upaya	274.533	59	maka semakin besar pula	226.165
7	berarti daya upaya yang	274.533	60	semakin besar pula usaha	246.996
8	daya upaya yang mendorong	198.098	61	besar pula usaha yang	301.755
9	upaya yang mendorong seseorang	234.276	62	pula usaha yang dilakukannya	226.888
10	yang mendorong seseorang untuk	233.879	63	usaha yang dilakukannya untuk	237.092
11	mendorong seseorang untuk melakukan	181.244	64	yang dilakukannya untuk mencapai	233.831
12	seseorang untuk melakukan sesuatu	211.318	65	dilakukannya untuk mencapai tujuan	209.635
13	untuk melakukan sesuatu dorongan	182.697	66	untuk mencapai tujuan motivasi	270.695
14	melakukan sesuatu dorongan dasar	116.177	67	mencapai tujuan motivasi sangat	295.682
15	sesuatu dorongan dasar yang	168.812	68	tujuan motivasi sangat diperlukan	226.734
16	dorongan dasar yang menggerakkan	175.775	69	motivasi sangat diperlukan dalam	184.175
17	dasar yang menggerakkan tingkah	179.425	70	sangat diperlukan dalam proses	182.024
18	yang menggerakkan tingkah laku	228.006	71	diperlukan dalam proses belajar	172.392
19	menggerakkan tingkah laku manusia	242.684	72	dalam proses belajar sebab	196.730
20	tingkah laku manusia disebut	287.165	73	proses belajar sebab seseorang	271.827

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No.	Hasil K-gram	Nilai Hash	No.	Hasil K-gram	Nilai Hash
21	laku manusia disebut sebagai	287.421	74	belajar sebab seseorang tidak	254.804
22	manusia disebut sebagai motivasi	301.082	75	sebab seseorang tidak akan	207.962
23	disebut sebagai motivasi motivasi	282.743	76	seseorang tidak akan melakukan	198.378
24	sebagai motivasi motivasi adalah	226.776	77	tidak akan melakukan aktivitas	131.696
25	motivasi motivasi adalah kekuatan	253.984	78	akan melakukan aktivitas belajar	170.503
26	motivasi adalah kekuatan yang	257.645	79	melakukan aktivitas belajar jika	182.337
27	adalah kekuatan yang muncul	213.194	80	aktivitas belajar jika tidak	212.137
28	kekuatan yang muncul dari	231.194	81	belajar jika tidak mempunyai	243.649
29	yang muncul dari dalam	153.604	82	jika tidak mempunyai motivasi	224.016
30	muncul dari dalam diri	126.667	83	tidak mempunyai motivasi dalam	176.297
31	dari dalam diri seseorang	183.121	84	mempunyai motivasi dalam belajar	215.104
32	dalam diri seseorang untuk	211.643	85	motivasi dalam belajar motivasi	238.271
33	diri seseorang untuk mendapatkan	205.969	86	dalam belajar motivasi menimbulkan	183.379
34	seseorang untuk mendapatkan tujuan	208.710	87	belajar motivasi menimbulkan gairah	237.923
35	untuk mendapatkan tujuan yang	200.368	88	motivasi menimbulkan gairah semangat	164.252
36	mendapatkan tujuan yang diinginkan	149.974	89	menimbulkan gairah semangat dan	107.096
37	tujuan yang diinginkan dan	161.586	90	gairah semangat dan merasa	188.351
38	yang diinginkan dan berusaha	173.352	91	semangat dan merasa senang	184.127
39	diinginkan dan berusaha melakukan	120.717	92	dan merasa senang untuk	242.443
40	dan berusaha melakukan perubahan	140.167	93	merasa senang untuk belajar	319.232
41	berusaha melakukan perubahan tingkah	184.355	94	senang untuk belajar semakin	241.771

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No.	Hasil <i>K-gram</i>	Nilai <i>Hash</i>	No.	Hasil <i>K-gram</i>	Nilai <i>Hash</i>
42	melakukan perubahan tingkah laku	179.463	95	untuk belajar semakin besar	268.352
43	perubahan tingkah laku yang	232.098	96	belajar semakin besar motivasi	264.074
44	tingkah laku yang lebih	216.988	97	semakin besar motivasi seseorang	256.444
45	laku yang lebih baik	202.953	98	besar motivasi seseorang maka	295.192
46	yang lebih baik demi	192.799	99	motivasi seseorang maka semakin	230.287
47	lebih baik demi memenuhi	180.581	100	seseorang maka semakin besar	244.094
48	baik demi memenuhi keinginannya	231.321	101	maka semakin besar pula	226.165
49	demi memenuhi keinginannya semakin	207.226	102	semakin besar pula energy	192.946
50	memenuhi keinginannya semakin besar	244.848	103	besar pula energi yang	247.705
51	keinginannya semakin besar dorongan	236.787	104	pula energi yang dimilikinya	189.808
52	semakin besar dorongan dari	203.579	105	energi yang dimilikinya untuk	200.006
53	besar dorongan dari dalam	191.583	106	yang dimilikinya untuk belajar	265.208

Setelah semua nilai *hash* dihitung dari dokumen 1, maka proses yang dilakukan selanjutnya adalah menghitung nilai *hash* dari dokumen 2, semua proses yang telah diterapkan pada dokumen 1 sama halnya dengan dokumen 2. Setelah itu proses selanjutnya membandingkan nilai *hash* dari kedua dokumen tersebut. Maka akan dilakukan pemilihan nilai *hash* yang sama dari kedua dokumen. Proses berikutnya yaitu menghitung persentase *similarity* dari kedua dokumen dengan *Cosine Similarity*.

4.1.3 Proses Ekstraksi Nilai *Hashing* Dokumen 2

Pada tahap ini dilakukan proses penginputan dokumen 2 kedalam sistem pemeriksaan kesamaan dokumen, dimana pada tahapan ini segala proses yang

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dilakukan sama dengan tahapan-tahapan pada saat penginputan dokumen 1 sebelumnya.

Pada tahapan ini dokumen 2 akan melalui beberapa tahapan yaitu, tahapan *preprocessing* yang terdiri dari *case folding*, *filtering*, dan *tokenizing*, hingga pencarian nilai *hash* pada setiap kata yang terdapat pada dokumen. Setelah dilakukan beberapa tahapan *Rabin-Karp* tersebut, maka sistem akan memeriksa nilai *similarity* (kesamaan) dari dua dokumen tersebut dengan menggunakan *Cosine Similarity*. Berikut ini merupakan contoh isi teks dari dokumen 2.

Tabel 4.6 Contoh Dokumen 2

Pengertian Motivasi Menurut Para Ahli – Motivasi berasal dari kata “motif” yang diartikan sebagai daya upaya yang mendorong seseorang untuk melakukan sesuatu. Menurut Sardiman (2006:73) motif merupakan daya penggerak dari dalam untuk melakukan kegiatan untuk mencapai tujuan.

Dalam proses belajar, motivasi sangat diperlukan, sebab seseorang yang tidak mempunyai motivasi dalam belajar, tidak akan mungkin melaksanakan aktivitas belajar. Motivasi diperlukan dalam menentukan intensitas usaha belajar bagi para siswa.

1. Tahapan *Preprocessing*

Berdasarkan pada *flowchart* 4.4, tahapan *preprocessing* yang harus dilalui oleh dokumen 2 yaitu *case folding*, *filtering*, dan *tokenizing*. Berikut contoh proses *preprocessing* pada dokumen 2.

a. *Case folding*

Case folding merupakan proses mengubah semua huruf kapital menjadi huruf kecil pada dokumen. Berikut hasil dari *case folding* pada dokumen 2.

Tabel 4.7 Dokumen 2 Setelah Melalui Proses *Case Folding*

pengertian motivasi menurut para ahli – motivasi berasal dari kata “motif” yang diartikan sebagai daya upaya yang mendorong seseorang untuk melakukan sesuatu. menurut sardiman (2006:73) motif merupakan daya penggerak dari dalam untuk melakukan kegiatan untuk mencapai tujuan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dalam proses belajar, motivasi sangat diperlukan, sebab seseorang yang tidak mempunyai motivasi dalam belajar, tidak akan mungkin melaksanakan aktivitas belajar. motivasi diperlukan dalam menentukan intensitas usaha belajar bagi para siswa.

b. *Filtering*

filtering merupakan proses penghilangan tanda baca yang kurang penting dengan menggunakan algoritme *stop list* (membuang tanda baca yang kurang penting). Berikut hasil dari *filtering* pada dokumen 2.

Tabel 4.8 Dokumen 2 Setelah Melalui Proses *Filtering* (*Stoplist Removal*)

pengertian motivasi menurut para ahli motivasi berasal dari kata motif yang diartikan sebagai daya upaya yang mendorong seseorang untuk melakukan sesuatu menurut sardiman 200673 motif merupakan daya penggerak dari dalam untuk melakukan kegiatan untuk mencapai tujuan dalam proses belajar motivasi sangat diperlukan sebab seseorang yang tidak mempunyai motivasi dalam belajar tidak akan mungkin melaksanakan aktivitas belajar motivasi diperlukan dalam menentukan intensitas usaha belajar bagi para siswa

c. *Tokenizing* (Proses Pembentukan *K-gram*)

Pembentukan *K-gram* merupakan proses *tokenisasi* atau pemotongan kata guna untuk membentuk pola kata, dimana pola katanya dalam bentuk *gram* dengan panjang *k*. Berikut hasil dari *tokenisasi* pada dokumen 2 dengan panjang

$$K = 4$$

Hasil *K-gram* dokumen 2

Tabel 4.9 Hasil *K-Gram* Dokumen 2

No.	Hasil <i>K-gram</i>	No.	Hasil <i>K-gram</i>
1	pengertian motivasi menurut para	33	kegiatan untuk mencapai tujuan
2	motivasi menurut para ahli	34	untuk mencapai tujuan dalam
3	menurut para ahli motivasi	35	mencapai tujuan dalam proses
4	para ahli motivasi berasal	36	tujuan dalam proses belajar
5	ahli motivasi berasal dari	37	dalam proses belajar motivasi
6	motivasi berasal dari kata	38	proses belajar motivasi sangat
7	berasal dari kata motif	39	belajar motivasi sangat diperlukan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

No.	Hasil K-gram	No.	Hasil K-gram
8	dari kata motif yang	40	motivasi sangat diperlukan sebab
9	kata motif yang diartikan	41	sangat diperlukan sebab seseorang
10	motif yang diartikan sebagai	42	diperlukan sebab seseorang yang
11	yang diartikan sebagai daya	43	sebab seseorang yang tidak
12	diartikan sebagai daya upaya	44	seseorang yang tidak mempunyai
13	sebagai daya upaya yang	45	yang tidak mempunyai motivasi
14	daya upaya yang mendorong	46	tidak mempunyai motivasi dalam
15	upaya yang mendorong	47	mempunyai motivasi dalam
16	yang mendorong seseorang	48	belajar motivasi dalam belajar tidak
17	untuk	49	dalam belajar tidak akan
18	mendorong seseorang untuk	50	belajar tidak akan mungkin
19	melakukan	51	belajar tidak akan mungkin
20	seseorang untuk melakukan sesuatu	52	melaksanakan
21	untuk melakukan sesuatu menurut	53	akan mungkin melaksanakan
22	melakukan sesuatu menurut sardiman	54	aktivitas
23	sesuatu menurut sardiman 200673	55	mungkin melaksanakan aktivitas
24	menurut sardiman 200673 motif	56	belajar melaksanakan aktivitas
25	sardiman 200673 motif merupakan	57	motivasi belajar motivasi
26	200673 motif merupakan daya	58	aktivitas belajar motivasi
27	motif merupakan daya penggerak	59	diperlukan
28	merupakan daya penggerak dari	60	belajar motivasi diperlukan dalam
29	daya penggerak dari dalam	61	motivasi diperlukan dalam
30	penggerak dari dalam untuk	62	menentukan
31	dari dalam untuk melakukan	63	diperlukan dalam menentukan
32	dalam untuk melakukan kegiatan		intensitas
	untuk melakukan kegiatan untuk		dalam menentukan intensitas
	melakukan kegiatan untuk mencapai		usaha
			menentukan intensitas usaha
			belajar
			intensitas usaha belajar bagi
			usaha belajar bagi para
			belajar bagi para siswa

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Proses Perhitungan Nilai Hash

Berdasarkan *flowchart* 4.8, berikut merupakan perhitungan nilai *hash* dari dokumen 2 pada hasil *K-gram* pertama dengan $k = 4$, basis $b = 11$, dan modulo $q = 100007$, pada kalimat “pengertian motivasi menurut para” adalah sebagai berikut.

Perhitungan untuk Kalimat 1:

$$\text{Hash}(w[0 \dots m-1]) = (w[0] * b^{m-1} + w[1] * b^{m-2} + \dots + w[m-1] * b^0) \bmod q$$

$$\begin{aligned} H_{(\text{pengertian})} &= (\text{ascii}(\text{p}) * 11^{(9)} + \text{ascii}(\text{e}) * 11^{(8)} + \text{ascii}(\text{n}) * 11^{(7)} + \text{ascii}(\text{g}) * 11^{(6)} \\ &+ \text{ascii}(\text{e}) * 11^{(5)} + \text{ascii}(\text{r}) * 11^{(4)} + \text{ascii}(\text{t}) * 11^{(3)} + \text{ascii}(\text{i}) * 11^{(2)} \\ &+ \text{ascii}(\text{a}) * 11^{(1)} + \text{ascii}(\text{n}) * 11^{(0)}) \bmod 100007 \end{aligned}$$

$$= (112 * 2.357.947.691 + 101 * 214.358.881 + 110 * 19.487.171 + 103 * 1.771.561 + 101 * 161.051 + 114 * 14.641 + 116 * 1.331 + 105 * 121 + 97 * 11 + 110 * 1) \bmod 100007$$

$$= (264.090.141.392 + 21.650.246.981 + 2.143.588.810 + 182.470.783 + 16.266.151 + 1.669.074 + 154.396 + 12.705 + 1.067 + 110) \bmod 100007$$

$$= 288.084.551.469 \bmod 100007$$

$$= 86.968$$

$$\begin{aligned} H_{(\text{motivasi})} &= (\text{ascii}(\text{m}) * 11^{(7)} + \text{ascii}(\text{o}) * 11^{(6)} + \text{ascii}(\text{t}) * 11^{(5)} + \text{ascii}(\text{i}) * 11^{(4)} \\ &+ \text{ascii}(\text{v}) * 11^{(3)} + \text{ascii}(\text{a}) * 11^{(2)} + \text{ascii}(\text{s}) * 11^{(1)} + \text{ascii}(\text{i}) * 11^{(0)}) \\ &\bmod 100007 \end{aligned}$$

$$= (109 * 19.487.171 + 111 * 1.771.561 + 116 * 161.051 + 105 * 14.641 + 118 * 1.331 + 97 * 121 + 115 * 11 + 105 * 1) \bmod 100007$$

$$= (2.124.101.639 + 196.643.271 + 18.681.916 + 1.537.305 + 157.058 + 11.737 + 1.265 + 105) \bmod 100007$$

$$= 2.341.134.296 \bmod 100007$$

$$= 70.433$$

$$\begin{aligned} H_{(\text{menurut})} &= (\text{ascii}(\text{m}) * 11^{(6)} + \text{ascii}(\text{e}) * 11^{(5)} + \text{ascii}(\text{n}) * 11^{(4)} + \text{ascii}(\text{u}) * 11^{(3)} \\ &+ \text{ascii}(\text{r}) * 11^{(2)} + \text{ascii}(\text{u}) * 11^{(1)} + \text{ascii}(\text{t}) * 11^{(0)}) \bmod 100007 \end{aligned}$$

$$= (109 * 1.771.561 + 101 * 161.051 + 110 * 14.641 + 117 * 1.331 + 114 * 121 + 117 * 11 + 116 * 1) \bmod 100007$$

$$= (193.100.149 + 16.266.151 + 1.610.510 + 155.727 + 13.794 + 1.287 + 116) \bmod 100007$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

$$\begin{aligned}
 &= 211.147.734 \bmod 100007 \\
 &= 32.957 \\
 H_{(\text{para})} &= (\text{ascii}(p) * 11^{(3)} + \text{ascii}(a) * 11^{(2)} + \text{ascii}(r) * 11^{(1)} + \text{ascii}(a) * 11^{(0)}) \\
 &\quad \bmod 100007 \\
 &= (112 * 1.331 + 97 * 121 + 114 * 11 + 97 * 1) \bmod 100007 \\
 &= (149.072 + 11.737 + 1.254 + 97) \bmod 100007 \\
 &= 162.160 \bmod 100007 \\
 &= 62.153
 \end{aligned}$$

$$\begin{aligned}
 Hash_1 &= 86.968 + 70.433 + 32.957 + 62.153 \\
 &= \mathbf{252.511}
 \end{aligned}$$

Perhitungan nilai *hash* dari hasil *K-gram* selanjutnya dapat dilihat dari tabel

4.10 yang dirujuk pada lampiran B yang merupakan contoh perhitungan manual nilai *hash* pada token-token berikutnya.

Tabel 4.10 Nilai Hash Setiap K-Gram Dokumen 2

No.	Hasil K-gram	Nilai Hash	No	Hasil K-gram	Nilai Hash
1	pengertian motivasi menurut para	252.511	33	kegiatan untuk mencapai tujuan	265.351
2	motivasi menurut para ahli	208.520	34	untuk mencapai tujuan dalam	207.601
3	menurut para ahli motivasi	208.520	35	mencapai tujuan dalam proses	201.172
4	para ahli motivasi berasal	254.551	36	tujuan dalam proses belajar	215.585
5	ahli motivasi berasal dari	238.587	37	dalam proses belajar motivasi	236.120
6	motivasi berasal dari kata	251.130	38	proses belajar motivasi sangat	328.479
7	berasal dari kata motif	239.481	39	belajar motivasi sangat diperlukan	266.902
8	dari kata motif yang	234.587	40	motivasi sangat diperlukan sebab	207.879
9	kata motif yang diartikan	225.851	41	sangat diperlukan sebab seseorang	219.882
10	motif yang diartikan sebagai	228.052	42	diperlukan sebab seseorang yang	194.278
11	yang diartikan sebagai daya	215.526	43	sebab seseorang yang tidak	238.832

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No.	Hasil K-gram	Nilai Hash	No	Hasil K-gram	Nilai Hash
12	diartikan sebagai daya upaya	216.540	44	seseorang yang tidak mempunyai	255.055
13	sebagai daya upaya yang	253.181	45	yang tidak mempunyai motivasi	243.052
14	daya upaya yang mendorong	198.098	46	tidak mempunyai motivasi dalam	176.297
15	upaya yang mendorong seseorang	234.276	47	mempunyai motivasi dalam belajar	215.104
16	yang mendorong seseorang untuk	233.879	48	motivasi dalam belajar tidak	219.097
17	mendorong seseorang untuk melakukan	181.244	49	dalam belajar tidak akan	191.888
18	seseorang untuk melakukan sesuatu	211.318	50	belajar tidak akan mungkin	274.528
19	untuk melakukan sesuatu. Menurut	161.839	51	tidak akan mungkin melaksanakan	278.268
20	melakukan sesuatu. menurut sardiman	160.963	52	akan mungkin melaksanakan aktivitas	242.763
21	sesuatu. menurut sardiman 200673	165.284	53	mungkin melaksanakan aktivitas belajar	289.605
22	menurut sardiman 200673 motif	191.356	54	melaksanakan aktivitas belajar motivasi	270.059
23	sardiman 200673 motif merupakan	175.856	55	aktivitas belajar motivasi diperlukan	182.958
24	200673 motif merupakan daya	148.279	56	belajar motivasi diperlukan dalam	174.543
25	motif merupakan daya penggerak	205.695	57	motivasi diperlukan dalam menentukan	160.842
26	merupakan daya penggerak dari	193.100	58	diperlukan dalam menentukan intensitas	116.403
27	daya penggerak dari dalam	182.982	59	dalam menentukan intensitas usaha	188.612
28	penggerak dari dalam untuk	211.435	60	menentukan intensitas usaha belajar	271.339
29	dari dalam untuk melakukan	149.698	61	intensitas usaha belajar bagi	238.380
30	dalam untuk melakukan kegiatan	168.598	62	usaha belajar bagi para	274.539

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No.	Hasil K-gram	Nilai Hash	No	Hasil K-gram	Nilai Hash
31	untuk melakukan kegiatan untuk	235.970	63	belajar bagi para siswa	234.290
32	melakukan kegiatan untuk mencapai	236.912			

4.1.4 Proses Pencarian Nilai Hash yang Sama

Setelah dilakukan berbagai macam tahapan, dan telah mendapatkan nilai *hash* dari dokumen 1 dan dokumen 2, maka proses yang akan dilakukan yaitu mencari nilai *hash* yang sama yang terdapat pada dokumen 1 dan dokumen 2. Berdasarkan *hashing* yang telah didapatkan dari dua dokumen diatas, maka *hashing* yang sama dari dokumen-dokumen tersebut sebagai berikut.

Tabel 4.11 Hashing yang Sama Pada Dokumen 1 Dan Dokumen 2

No	Hashing yang Sama Pada Dokumen 1		Hashing yang Sama Pada Dokumen 2	
1	motivasi berasal dari kata	251.130	motivasi berasal dari kata	251.130
2	berasal dari kata motif	239.481	berasal dari kata motif	239.481
3	dari kata motif yang	234.587	dari kata motif yang	234.587
4	daya upaya yang mendorong	198.098	daya upaya yang mendorong	198.098
5	upaya yang mendorong seseorang	234.276	upaya yang mendorong seseorang	234.276
6	yang mendorong seseorang untuk	233.879	yang mendorong seseorang untuk	233.879
7	mendorong seseorang untuk melakukan	181.244	mendorong seseorang untuk melakukan	181.244
8	seseorang untuk melakukan sesuatu	211.318	seseorang untuk melakukan sesuatu	211.318
9	tidak mempunyai motivasi dalam	176.297	tidak mempunyai motivasi dalam	176.297
10	mempunyai motivasi dalam belajar	215.104	mempunyai motivasi dalam belajar	215.104

Jumlah *hash* yang sama pada dokumen 1 dan dokumen 2 adalah:

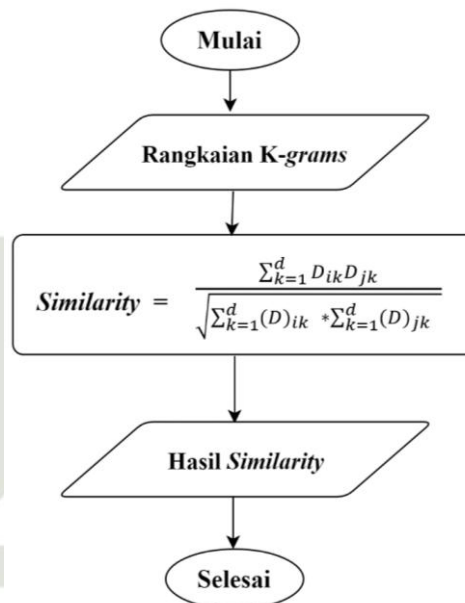
$$\sum_{k=1}^d D_{ik} D_{jk} = 10 \text{ hash value}$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.1.5 Proses Perhitungan Persentase Kesamaan (*Similarity*)

Tahapan-tahapan perhitungan kesamaan dokumen akan digambarkan pada *flowchart* berikut ini.



Gambar 4.7 Flowchart Proses Pemeriksaan Kesamaan

Rangkaian tahapan telah dilakukan hingga mendapatkan beberapa nilai *hash* yang sama dari kedua dokumen, dengan jumlah *hash* pada dokumen 1 dapat dilihat pada tabel 4.4 sebanyak **106**, jumlah *hash* dokumen 2 pada tabel 4.9 sebanyak **63**, dan *hash* yang sama pada kedua dokumen pada tabel 4.11 sebanyak **10**. Proses selanjutnya menghitung *similarity* atau persentase kesamaan antar dua dokumen dengan menggunakan persamaan (2.3).

Hasil *similarity* didapat dari total jumlah *hash* yang sama dibagi dengan akar dari total *hash* pada dokumen 1 yang dikalikan dengan akar dari total *hash* pada dokumen 2 dan dikali 100%. Berikut ini adalah proses menghitung *similarity* dua dokumen tersebut:

$$\begin{aligned}
 \text{Sim } (D_i, D_j) &= \frac{\sum_{k=1}^d D_{ik} D_{jk}}{\sqrt{\sum_{k=1}^d (D)_{ik} * \sum_{k=1}^d (D)_{jk}}} \times 100\% \\
 &= \frac{10}{\sqrt{106} \sqrt{63}} \times 100\%
 \end{aligned}$$

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

$$= \frac{10}{81,71} \times 100\%$$

$$= 12,24 \%$$

Hasil perhitungan *similarity* dari kedua dokumen tersebut adalah sebesar 12,24%.

Berikut merupakan hasil *highlight* kata yang sama dari dokumen 1 dan dokumen 2

Tabel 4.12 Hasil *Highlight* Kata Yang Sama Kedua Dokumen

Dokumen 1	Dokumen 2
<p>motivasi berasal dari kata “motif” yang berarti daya upaya yang mendorong seseorang untuk melakukan sesuatu dorongan dasar yang menggerakkan tingkah laku manusia disebut sebagai motivasi motivasi adalah kekuatan yang muncul dari dalam diri seseorang untuk mendapatkan tujuan yang diinginkan dan berusaha melakukan perubahan tingkah laku yang lebih baik demi memenuhi keinginannya semakin besar dorongan dari dalam diri seseorang maka semakin besar pula usaha yang dilakukannya untuk mencapai tujuan motivasi sangat diperlukan dalam proses belajar sebab seseorang tidak akan melakukan aktivitas belajar jika tidak mempunyai motivasi dalam belajar motivasi menimbulkan gairah, semangat dan merasa senang untuk belajar. semakin besar motivasi seseorang maka semakin besar pula energi yang dimilikinya untuk belajar</p>	<p>pengertian motivasi menurut para ahli motivasi berasal dari kata “motif” yang diartikan sebagai daya upaya yang mendorong seseorang untuk melakukan sesuatu. menurut sardiman 200673 motif merupakan daya penggerak dari dalam untuk melakukan kegiatan untuk mencapai tujuan dalam proses belajar motivasi sangat diperlukan sebab seseorang yang tidak mempunyai motivasi dalam belajar tidak akan mungkin melaksanakan aktivitas belajar motivasi diperlukan dalam menentukan intensitas usaha belajar bagi para siswa</p>

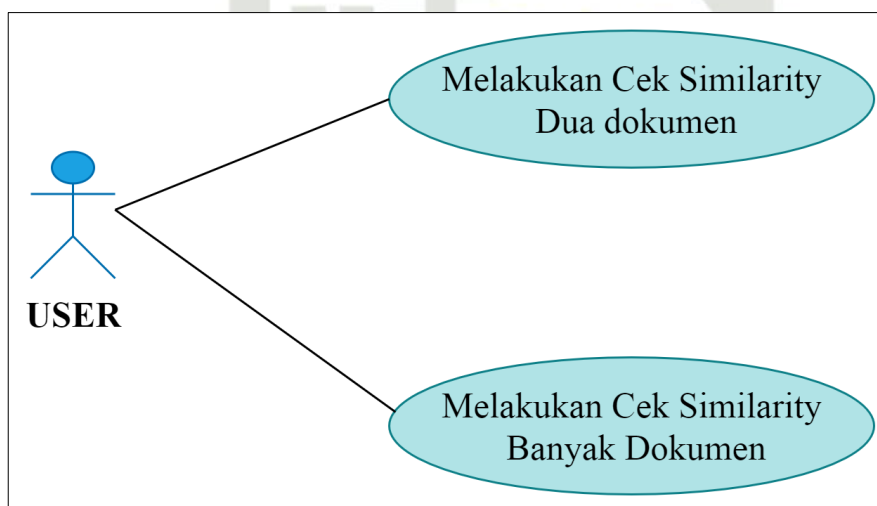
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.1.6 Analisa Fungsional Sistem

Analisa fungsional sistem untuk pemeriksaan kesamaan isi terhadap dokumen terdiri dari *Use case Diagram*, *Use case Specification*, *Class Diagram*, dan *Sequence Diagram*.

4.1.6.1 *Use case Diagram*

Use case diagram merupakan sebuah diagram yang menggambarkan interaksi pengguna dengan sistem, dengan kata lain pada *use case diagram* ini berguna untuk mengetahui fungsi-fungsi yang ada dalam sebuah sistem dan siapa yang berhak menjalankan setiap fungsi-fungsi tertentu. Gambar 4.8 berikut ini adalah *use case diagram* yang menjelaskan fungsi-fungsi yang ada dalam sebuah sistem dan siapa saja yang berhak menjalankan setiap fungsi-fungsi tertentu dari dalam pemeriksaan kesamaan isi dokumen.



Gambar 4.8 *Use case Diagram*

Pada *use case* diatas terdapat seorang *actor* yaitu *user* yang dapat mengakses sistem pemeriksaan kesamaan dokumen. User dapat melakukan pemeriksaan dua dokumen dan banyak dokumen. Saat *user* melakukan pemeriksaan dua dokumen, *user* dapat menginputkan dokumen 1 dan dokumen 2, memasukkan nilai *K-gram* untuk mendapatkan hasil persentase kesamaan dari dokumen-dokumen tersebut.

Selain proses untuk dua dokumen, *user* juga dapat melakukan proses pemeriksaan banyak dokumen. Pada fungsi *use case* ini tidak berbeda jauh dengan *use case* pemeriksaan dua dokumen, *user* dapat menginputkan beberapa dokumen,

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

atau memasukkan nilai *K-gram* yang diinginkan untuk melakukan pemeriksaan kesamaan pada dokumen-dokumen yang telah diinput tersebut.

4.1.6.2 Use case Spesification

Pada *use case spesification* ini setiap *use case* yang terdapat pada *use case diagram* harus dijelaskan alur kerjanya dengan menggunakan skenario. Berikut adalah penjabaran dari *use case spesification* berdasarkan *use case diagram* sistem pemeriksaan kesamaan dokumen pada tabel.

1. Use Case Spesification Melakukan Cek Similarity Dua Dokumen

Use case spesification melakukan cek *similarity* dua dokumen menjelaskan skenario dari *use case* melakukan cek *similarity* dua dokumen. Proses ini merupakan proses ketika *user* sudah menekan tombol *button* “Pemeriksaan Dua Dokumen” pada halaman utama sistem.

Tabel 4.13 Use Case Spesification Melakukan Cek Similarity Dua Dokumen

Aktor Utama	User
Kondisi Awal	Dokumen dan nilai K harus sudah terisi
Kondisi Akhir	Pemeriksaan kesamaan dua dokumen berhasil dilakukan dengan menampilkan hasil dari proses yang dijalankan oleh sistem.
Main Succes Scenario	<ol style="list-style-type: none"> 1. <i>Usecase</i> dimulai ketika <i>user</i> telah memilih dua dokumen dan memasukkan nilai <i>K</i>. 2. <i>User</i> menekan tombol proses 3. Sistem melakukan pembacaan dokumen kedalam bentuk teks. 4. Sistem melakukan proses <i>case folding</i> berdasarkan teks yang dihasilkan dari pembacaan dokumen. 5. Sistem melakukan proses <i>filtering (stoplist removal)</i> berdasarkan hasil <i>case folding</i>. 6. Sistem melakukan proses <i>tokenizing</i> berdasarkan hasil <i>case folding</i> dan <i>filtering</i>. 7. Sistem menghitung nilai <i>hash</i> berdasarkan hasil <i>tokenizing</i>. 8. Sistem mencari nilai <i>hash</i> yang sama antar dua dokumen 9. Sistem menghitung nilai persentase kesamaan dokumen menggunakan <i>Cosine Similarity</i> berdasarkan nilai <i>hash</i> yang telah dihasilkan dari dua dokumen. 10. Sistem menampilkan hasil perbandingan dua teks dengan menandai token yang sama.
Alternative Scenario	Jika dokumen belum diinputkan, maka proses pemeriksaan tidak bisa dilakukan dan menampilkan pesan “Silahkan Pilih Dokumen Dahulu”

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Use Case Spesification Melakukan Cek *similarity* Banyak Dokumen

Use Case spesification melakukan cek *similarity* banyak dokumen merupakan penjelasan skenario dari *use case* melakukan cek *similarity* banyak dokumen yang dilakukan oleh *user* ketika ingin melakukan pemeriksaan dengan jumlah dokumen lebih dari dua dokumen.

Tabel 4.14 Use Case Spesification Melakukan Cek *similarity* Banyak Dokumen

Aktor Utama	<i>User</i>
Kondisi Awal	Dokumen dan nilai <i>K</i> harus sudah terisi
Kondisi Akhir	Pemeriksaan kesamaan banyak dokumen berhasil dilakukan dengan menampilkan hasil dari proses yang dijalankan oleh sistem.
Main Succes Scenario	<ol style="list-style-type: none"> 1. <i>Usecase</i> dimulai ketika <i>user</i> telah memilih beberapa dokumen dan memasukkan nilai <i>K</i>. 2. <i>User</i> menekan tombol proses 3. Sistem melakukan pembacaan dokumen kedalam bentuk teks. 4. Sistem melakukan proses <i>case folding</i> berdasarkan teks yang dihasilkan dari pembacaan dokumen. 5. Sistem melakukan proses <i>filtering (stoplist removal)</i> berdasarkan hasil <i>case folding</i>. 6. Sistem melakukan proses <i>tokenizing</i> berdasarkan hasil <i>case folding</i>. 7. Sistem menghitung nilai <i>hash</i> berdasarkan hasil <i>tokenizing</i>. 8. Sistem mencari nilai <i>hash</i> yang sama antar dokumen. 9. Sistem menghitung nilai persentase kesamaan antar dokumen menggunakan <i>Cosine Similarity</i> berdasarkan nilai <i>hash</i> yang telah dihasilkan dari beberapa dokumen. 10. Sistem menampilkan hasil perbandingan dua teks dengan menandai token yang sama.
Alternative Scenario	Jika dokumen belum diinputkan, maka proses pemeriksaan tidak bisa dilakukan dan menampilkan pesan “Silahkan Pilih Dokumen Dahulu”

4.1.6.3 Sequence Diagram

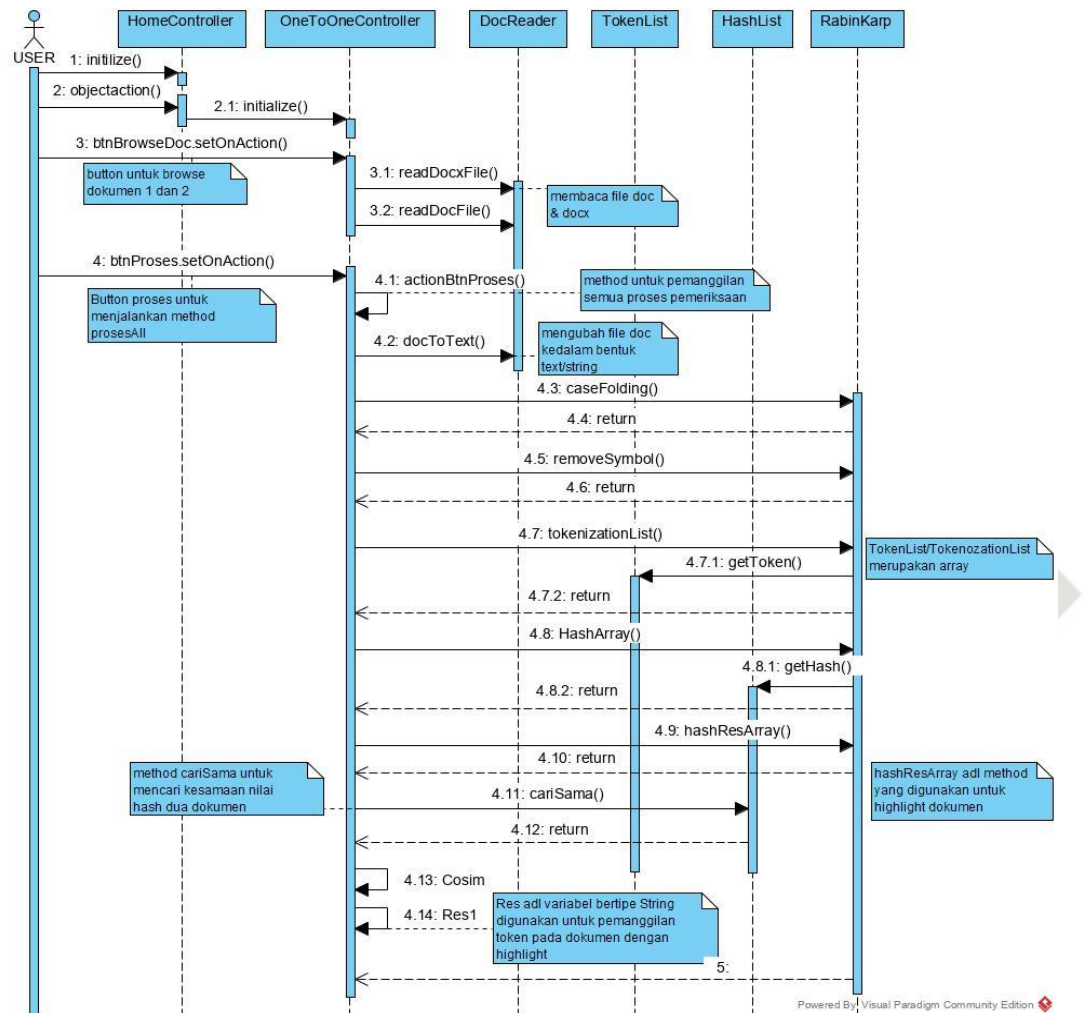
Sequence diagram merupakan sebuah diagram yang digunakan untuk menunjukkan interaksi *object* dan menggambarkan rangkaian pesan yang dikirim antara *object-object* tersebut pada saat interaksi. Semua pesan dideskripsikan dalam urutan pada eksekusi. Gambar-gambar berikut ini merupakan penggambaran interaksi antar *object* melalui *sequence diagram* dari setiap *use case* sistem pemeriksaan kesamaan isi dokumen.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Sequence Diagram Melakukan Cek Similarity Dua Dokumen

Gambar 4.9 berikut ini merupakan penggambaran interaksi antar *object* pada *sequence diagram* untuk melakukan cek *similarity* pada dua dokumen yang dilakukan oleh *user*.



Gambar 4.9 Sequence Diagram Melakukan Cek Similarity Dua Dokumen

Pada gambar diatas, terdapat beberapa *class*. *Class* pertama ketika menjalankan sistem yaitu *class HomeController*. Pada *class* berisi *method* `initiliaze()` dan `objectAction()` yang mengarah pada *class OneToOneController*. *Class OneToOneController* ini adalah *class* yang digunakan untuk melakukan pemeriksaan kesamaan dua dokumen. pada *class* ini terdapat beberapa aksi yang digunakan untuk pemeriksaan. Untuk aksi pemilihan dokumen pada *class OneToOneController* terdapat pada `btnBrowseDoc.setAction()`, dan `btnProses.setAction()` untuk tombol proses agar sistem menjalankan

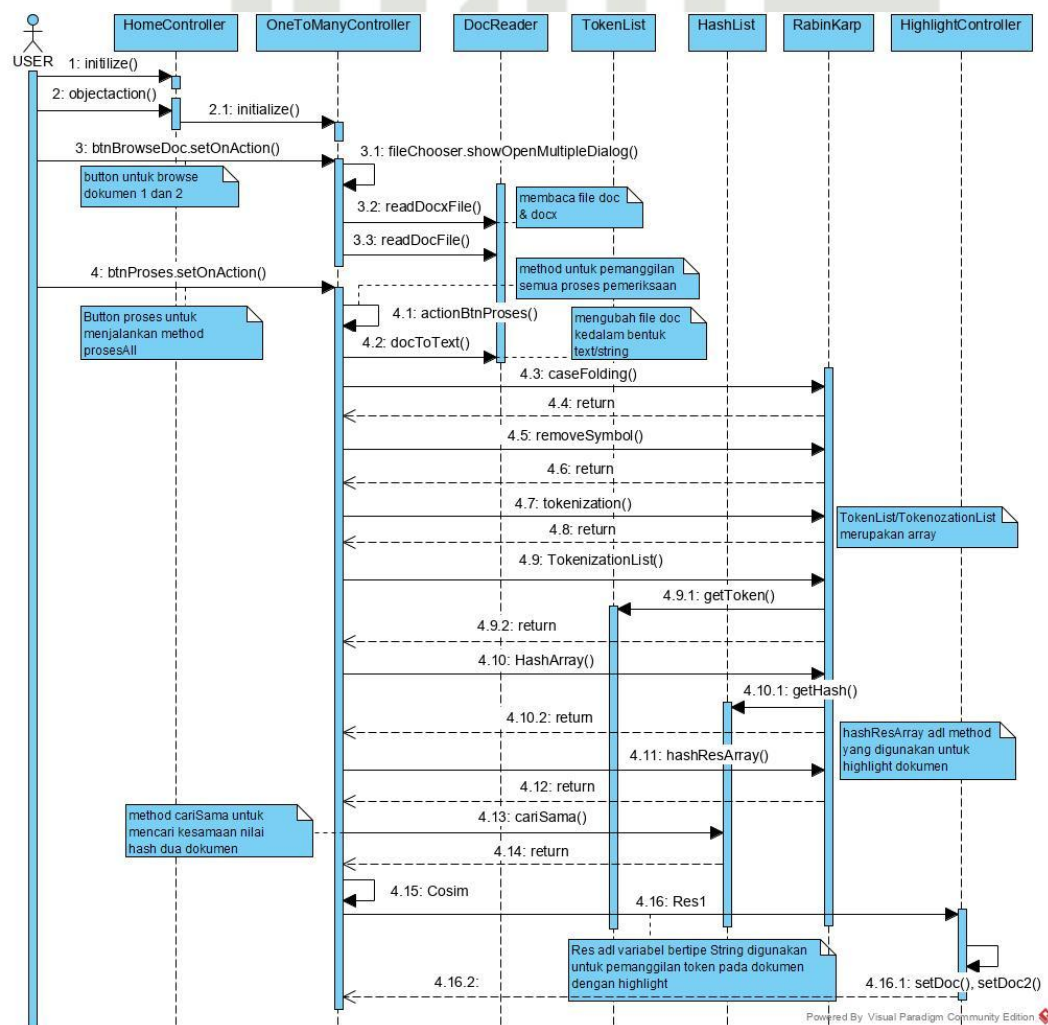
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

pemeriksaan dengan menggunakan algoritme *Rabin-Karp* dan *Cosine Similarity*. Pada class *OneToOneController* semua proses *Rabin-Karp* akan dipanggil dari *caseFolding*, *removalSymbol*, *tokenization*, *hashing*, hingga masuk pada persamaan *Cosine Similarity* untuk perhitungan persentase kesamaan antar Dokumen.

2. Sequence Diagram Melakukan Cek Similarity Banyak Dokumen

Gambar 4.10 berikut ini merupakan penggambaran interaksi antar *object* pada *sequence diagram* untuk melakukan cek *similarity* banyak dokumen yang dilakukan oleh *user*.



Gambar 4.10 Sequence Diagram Melakukan Cek Similarity Banyak Dokumen

Pada gambar diatas, terdapat beberapa *class*. *Class* pertama ketika menjalankan sistem yaitu *class HomeController*. Pada *class* berisi *method* *initiliaze()* dan

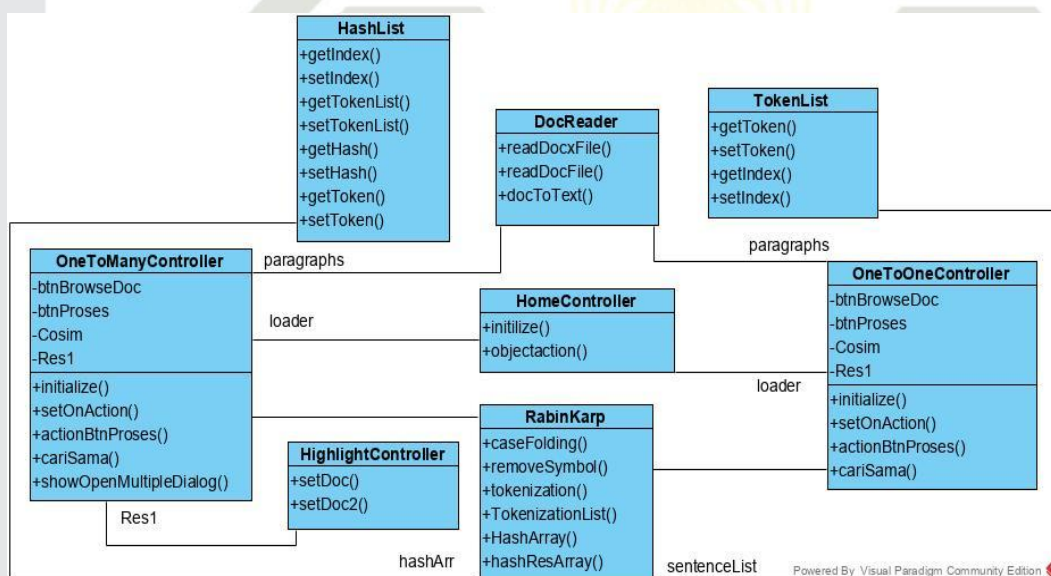
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

objectAction() yang mengarah pada *class OneToManyController*. *Class OneToManyController* ini adalah *class* yang digunakan untuk melakukan pemeriksaan kesamaan banyak dokumen. Setelah itu *user* melakukan aksi dengan menginputkan banyak dokumen dengan *method fileChooser.ShowOpenMultipleDialog()* dan melakukan proses untuk menjalankan pemeriksaan dengan algoritme-algoritme yang telah dibentuk. Pada *OneToManyController* semua proses *Rabin-Karp* akan dipanggil dari *caseFolding*, *removalSymbol*, *tokenization*, *hashing*, hingga *similarity*.

4.1.6.4 Class Diagram

Class diagram merupakan sebuah bentuk diagram yang menggambarkan struktur serta deskripsi dari *class* serta hubungan yang terjadi antar *class*. Pada sistem pemeriksaan kesamaan isi dokumen ini terdapat beberapa *class* untuk membangun sistem, *class-class* tersebut dapat dilihat pada gambar 4.11 berikut ini.



Gambar 4.11 Class Diagram

Pada *class diagram* diatas, sistem pemeriksaan kesamaan dokumen memiliki beberapa *class* diantaranya yaitu *class OneToOneController* untuk pemeriksaan kesamaan dua dokumen, *class OneToManyController* untuk pemeriksaan banyak dokumen, *class DocReader* untuk pembacaan dokumen teks yang berekstensi *.doc* dan *.docx*, *class HomeController* untuk halaman utama dari sistem, *class Hashlist* dan *Tokenlist* yang digunakan pada *class Rabin-Karp* untuk diproses dan mengirimkan hasilnya dalam bentuk *arraylist*, serta *class highlightController* untuk

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

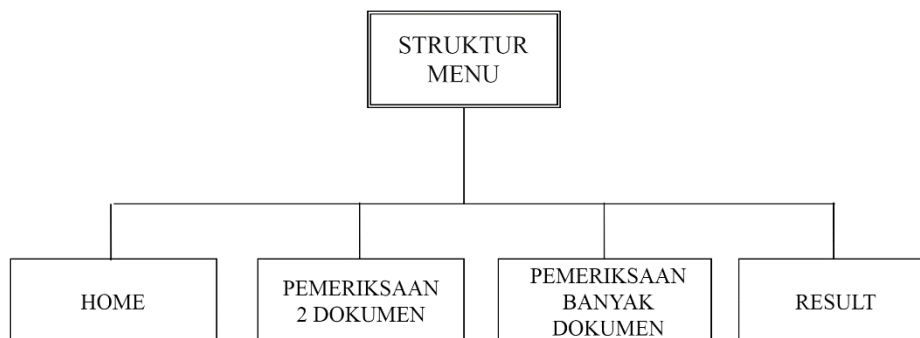
menampilkan hasil penanda kata-kata yang terdapat kesamaan pada dokumen-dokumen yang diuji. Pada setiap *class* ini terdapat *association* yaitu hubungan antar dua *class*. Contoh pada *class OneToOneController* dan *OneToManyController* terdapat asosiasi dengan nama *loader* yaitu sebagai nama variabel dari *instance HomeController*. Dan begitu juga variabel-variabel yang menghubungkan antar *class* lainnya.

4.2 Perancangan

Perancangan adalah suatu bentuk gambaran atau perencanaan, dan pembuatan sketsa yang sesuai dengan tujuan yang telah didapatkan dari hasil analisis sistem sebelumnya. Pada penelitian ini dilakukan perancangan terhadap struktur menu serta perancangan tampilan antarmuka pada sistem yang akan dibangun.

4.2.1 Perancangan Struktur Menu

Perancangan struktur menu adalah tahap merancang menu-menu yang dapat digunakan *user* untuk menjalankan sistem. Pada sistem pemeriksaan dokumen ini terdapat beberapa menu yaitu menu *Home*, menu pemeriksaan dua dokumen, menu pemeriksaan banyak dokumen, dan *result* untuk menampilkan *highlight* dokumen yang telah diuji.



Gambar 4.12 Struktur Menu Sistem

4.2.2 Perancangan Tampilan Antarmuka

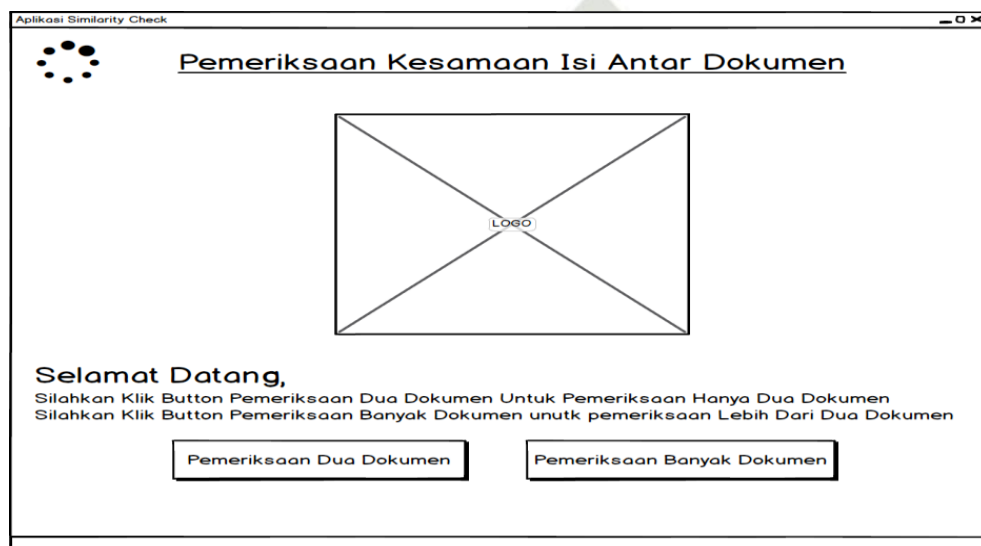
Antarmuka merupakan tampilan sistem yang dirancang yang berfungsi untuk memudahkan *user* dalam menggunakan sistem. Pada Sistem pemeriksaan kesamaan isi dokumen ini terdapat beberapa tampilan antarmuka, yaitu sebagai berikut:

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Tampilan *Home* Sistem

Halaman *Home* merupakan halaman yang pertama kali muncul ketika akan menjalankan sistem pemeriksaan kesamaan dokumen. Halaman *Home* berisi logo dari sistem, dan dua tombol *button* “Pemeriksaan Dua Dokumen” dan “Pemeriksaan Banyak Dokumen”. Berikut merupakan rancangan tampilan dari halaman *Home* pada gambar 4.13



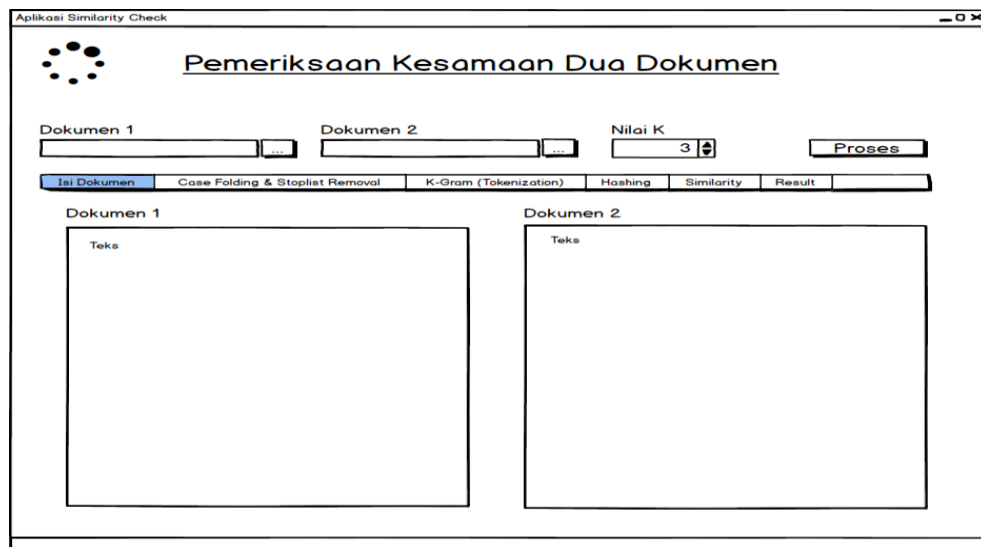
Gambar 4.13 Tampilan Menu *Home*

2. Tampilan Halaman Pemeriksaan Dua Dokumen

Halaman Pemeriksaan Dua Dokumen merupakan halaman untuk pemeriksaan kesamaan dua dokumen. Pada halaman ini terdapat *button browse* untuk memilih *file*, *spinner K-gram*, *button proses*, serta hasil dari pemrosesan dokumen yang terdiri dari *case folding*, *tokenizing*, *hashing*, *similarity*, serta *result*. Berikut merupakan rancangan tampilan dari halaman Pemeriksaan Dua Dokumen pada gambar 4.14.

Hak Cipta Diindungi Undang-Undang

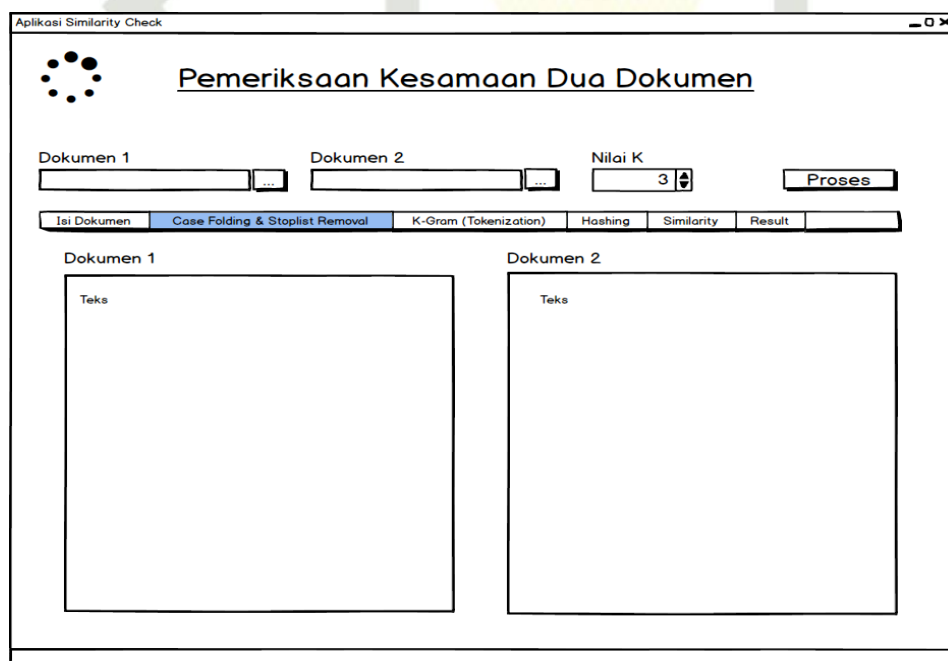
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4.14 Tampilan Menu Pemeriksaan Dua Dokumen

3. Tampilan Tab *Case folding & Stoplist*

Tampilan tab *case folding & stoplist* berisi isi dokumen yang telah melalui proses *case folding & stoplist*. Berikut merupakan rancangan tampilan dari tab *case folding & stoplist*. pada gambar 4.15.



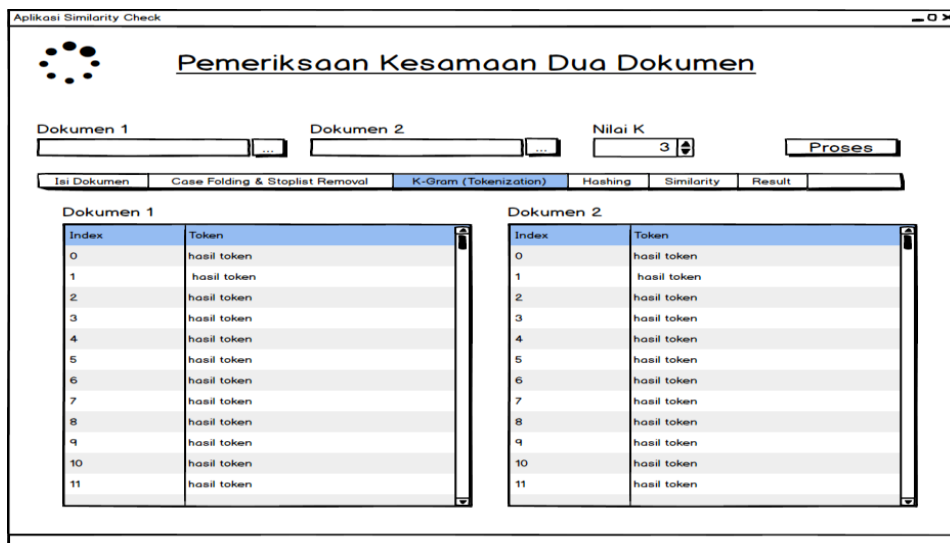
Gambar 4.15 Tampilan Tab *Case folding*

4. Tampilan Tab *Tokenization*

Tampilan tab *tokenizing* berisi hasil *token* dari dokumen yang diinput. Berikut merupakan rancangan tampilan dari tab *tokenizing* pada gambar 4.16.

Hak Cipta Dilindungi Undang-Undang

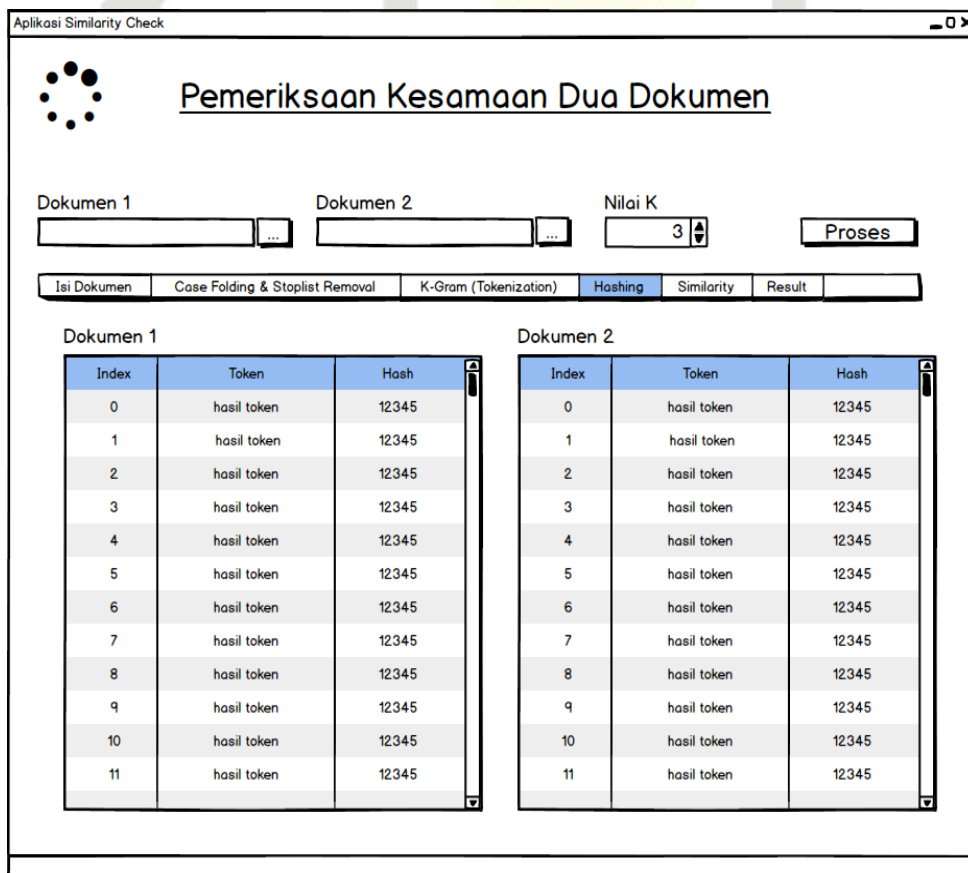
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4.16 Tampilan Tab *Tokenizing*

5. Tampilan Tab *Hashing*

Tampilan tab *hashing* berisi hasil *token* dokumen serta nilai *hash* dari *token* tersebut. Berikut merupakan rancangan tampilan dari tab *hashing* pada gambar 4.17.



Gambar 4.17 Tampilan Tab *Hashing*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

6. Tampilan Tab *Similarity*

Tampilan tab *similarity* berisi hasil *token* yang sama antara kedua dokumen serta nilai persentase *similarity* dari kedua dokumen. Berikut merupakan rancangan tampilan dari tab *similarity* pada gambar 4.18.

The screenshot shows the 'Similarity' tab of the 'Pemeriksaan Kesamaan Dua Dokumen' application. At the top, there are input fields for 'Dokumen 1', 'Dokumen 2', and 'Nilai K' (set to 3), along with a 'Proses' button. Below these is a navigation bar with tabs: 'Isi Dokumen', 'Case Folding & Stoplist Removal', 'K-Gram (Tokenization)', 'Hashing', 'Similarity' (active), and 'Result'. The main content area is divided into two sections. On the left, 'Daftar Hash yang Sama' contains a table with 12 rows, each showing an index, the token 'hasil token', and the hash '12345'. On the right, 'Hasil Persentase Similarity' displays '10.00%' in a large box.

Index	Token	Hash
0	hasil token	12345
1	hasil token	12345
2	hasil token	12345
3	hasil token	12345
4	hasil token	12345
5	hasil token	12345
6	hasil token	12345
7	hasil token	12345
8	hasil token	12345
9	hasil token	12345
10	hasil token	12345
11	hasil token	12345

Gambar 4.18 Tampilan Tab *Similarity*

7. Tampilan Tab *Result*

Tampilan tab *result* berisi hasil isi dokumen yang terdapat *highlight* pada kata yang sama setelah dilakukan pemeriksaan kesamaan. Berikut merupakan rancangan tampilan dari tab *result* pada gambar 4.19.

The screenshot shows the 'Result' tab of the 'Pemeriksaan Kesamaan Dua Dokumen' application. It features the same top navigation and input fields as the previous tab. The 'Result' tab is active. The main content area has two sections: 'Highlight Dokumen 1' and 'Highlight Dokumen 2'. Each section contains a text box with the document content, where matching words are highlighted in red. The text in both boxes is: 'sistem pemeriksaan kesamaan antar dokumen menggunakan algoritma rabin karp dan cosine similarity universitas sultan syarif kasim riau fakultas sains dan teknologi jurusan teknik informatika'.

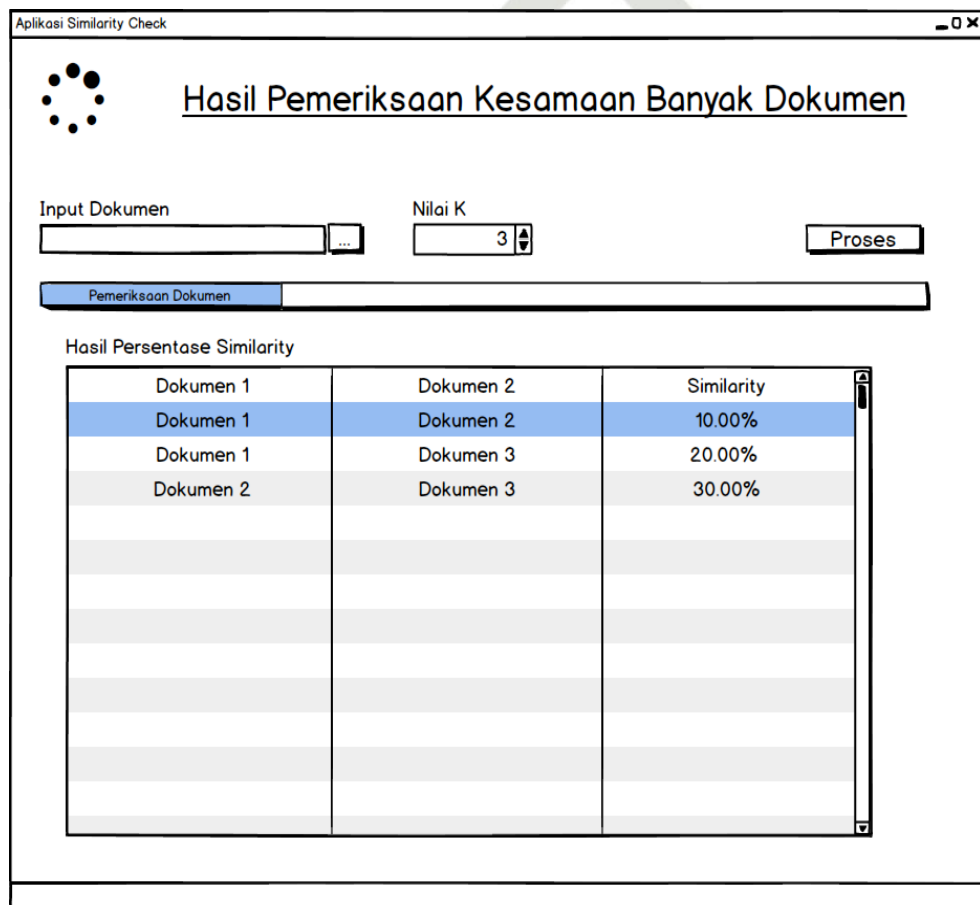
Gambar 4.19 Tampilan Tab *Result*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

8. Tampilan Halaman Pemeriksaan Banyak Dokumen

Halaman Pemeriksaan Banyak Dokumen merupakan halaman untuk pemeriksaan kesamaan lebih dari dua dokumen. Pada halaman ini terdapat *button browse* untuk memilih beberapa *file*, *spinner K-gram*, *button proses*, serta hasil persentase *similarity* dari pemrosesan dokumen. Berikut merupakan rancangan tampilan dari halaman Pemeriksaan Banyak Dokumen pada gambar 4.20.



The screenshot shows a web application window titled 'Aplikasi Similarity Check'. The main heading is 'Hasil Pemeriksaan Kesamaan Banyak Dokumen'. Below the heading, there is an 'Input Dokumen' field with a browse button, a 'Nilai K' spinner set to 3, and a 'Proses' button. A progress bar labeled 'Pemeriksaan Dokumen' is shown below. The results are displayed in a table titled 'Hasil Persentase Similarity'.

Dokumen 1	Dokumen 2	Similarity
Dokumen 1	Dokumen 2	10.00%
Dokumen 1	Dokumen 3	20.00%
Dokumen 2	Dokumen 3	30.00%

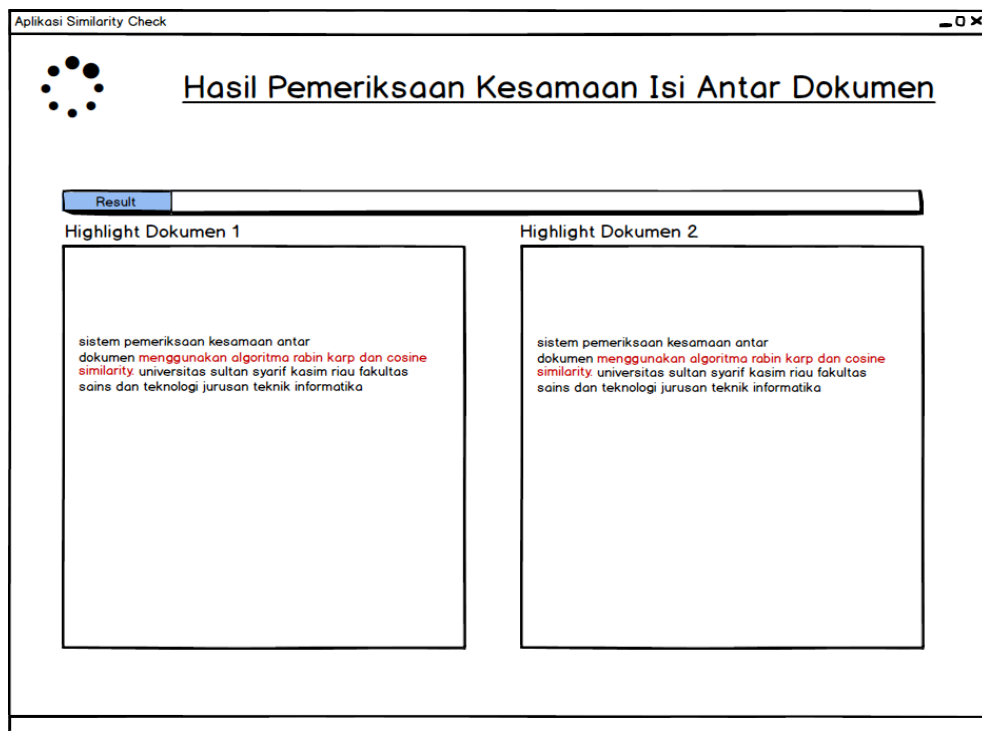
Gambar 4.20 Tampilan Halaman Pemeriksaan Banyak Dokumen

9. Tampilan Halaman *Result* dari Pemeriksaan Banyak Dokumen

Tampilan halaman *result* akan muncul jika salah satu hasil perbandingan dua dokumen yang berupa tabel *view* pada halaman Pemeriksaan Banyak Dokumen di klik. Halaman *result* terdiri dari dua dokumen yang terdapat *highlight* pada kata yang sama diantara dokumen tersebut. Berikut merupakan rancangan tampilan dari halaman *result* Pemeriksaan Banyak Dokumen pada gambar 4.21.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.



Gambar 4.21 Tampilan Halaman *Result* Pemeriksaan Banyak Dokumen

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB VI

PENUTUP

6.1 Kesimpulan

Kesimpulan yang dapat diambil berdasarkan pembahasan dari bab-bab sebelumnya pada penelitian pemeriksaan kesamaan dokumen ini yaitu :

1. Sistem pemeriksaan dokumen yang dirancang dan dibangun dengan menerapkan algoritme *Rabin-Karp* dan *Cosine Similarity* untuk pemeriksaan dokumen tugas makalah mahasiswa pada mata kuliah TTKI telah selesai dibangun dan memberikan hasil *output similarity* sesuai dengan tingkat kesamaan antar dokumen yang diuji.
2. Berdasarkan pengujian fungsional algoritme yang telah dilakukan, didapatkan bahwa hasil proses algoritme *Rabin-Karp* dan *Cosine Similarity* yang diterapkan pada sistem sama dengan hasil perhitungan manual dengan presentase *similarity* sebesar 12,24% pada dua dokumen yang diuji.
3. Berdasarkan pengujian tata letak kalimat yang telah dilakukan, didapatkan bahwa nilai *similarity* akan tetap sama walaupun letak kalimat pada isi dokumen telah diubah.
4. Berdasarkan pengujian kesamaan dokumen yang telah dilakukan, didapatkan bahwa sistem akan menghasilkan presentase *similarity* 100% apabila dokumen yang diuji merupakan dua dokumen yang sama. Sebaliknya, sistem akan menghasilkan presentase *similarity* 0% apabila dokumen yang diuji berisi kalimat yang berbeda dan kalimat yang telah diparafrase.
5. Berdasarkan pengujian parameter *K-gram* yang telah dilakukan, didapatkan bahwa semakin kecil nilai *K-gram* yang dipilih maka semakin tinggi hasil *similarity* yang diperoleh. Pada pengujian parameter *K-gram*, nilai *similarity* tertinggi terletak pada $K=3$ yaitu sebesar 18,54 %.
6. Berdasarkan pengujian banyak dokumen yang telah dilakukan, terdapat sebanyak 150 kombinasi dari 15 *file* dokumen TTKI yang diuji dengan hasil persentase nilai *similarity* yang berbeda-beda.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

7. Berdasarkan pengujian perbandingan antara sistem pemeriksaan kesamaan dokumen dengan *plagiarism checker x*, didapatkan bahwa hasil perbandingan kedua aplikasi ini menghasilkan *similarity* yang bervariasi dari 15 *file* yang diuji, *similarity* yang didapatkan pada pengujian dokumen 1 dan 2 sebesar 2,52% yang dihasilkan sistem pemeriksaan dokumen dan 1,16 % yang dihasilkan *plagiarism checker x*. Rata-rata selisih nilai *similarity* dari 15 *file* dokumen yang dihasilkan kedua aplikasi ini sebesar 0,499 %.

6.2 Saran

Terdapat beberapa hal yang dapat menjadi saran untuk penyempurnaan sistem ini agar lebih baik dimasa yang akan datang antara lain:

1. Sistem ini akan lebih baik jika dapat membaca *file* dokumen dalam format lain seperti *.pdf*, *.txt*, *.xls*, dan *.ppt*
2. Sistem akan lebih baik jika dapat dikembangkan dengan berbasis *web* agar dapat diakses melalui internet.
3. Penelitian dapat dilanjutkan dengan menggunakan algoritme pemeriksaan kesamaan yang berbeda untuk menerapkan algoritme yang lebih baik dalam pemeriksaan kesamaan dokumen.

DAFTAR PUSTAKA

- Ernawati, E., Anindito, & Atmojo, R. N. (2014). Sistem Pendeteksi Plagiarisme Untuk Tugas Akhir Mahasiswa Di Universitas Bina Nusantara Studi Pendahuluan. *Humaniora*.
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook Advanced*. New York: United States of America by Cambridge University Press.
- Hasan, E. G., Wicaksana, A., & Hansun, S. (2018). The Implemnetation of Winnowing Algorithm for Plagiarism Detection in Moodle-based E-learning. *IEEE Computer Society*, 322.
- J.Money, R. (2006). Machine Learning Text Categorization. *University of Texas at Austin*.
- Joshi, O. S., Upadhyay, B. r., & M, S. (2017). Parallelized Advanced Rabin-Karp Algorithm For String Matching. *ICCUBEA, IEEE*.
- K.Popat, M., Desmukh, M. B., & Metre, M. A. (2017). Hierarchical Document Clustering based on Cosine Similarity Measure. *IEEE*.
- Leonardo, B., & Hansun, S. (2017). Text Documents Plagiarism Detection Using Rabin-Karp and Jaro-Winkler Distance Algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*.
- Mahmood, Q., Qadir, M. A., & Afzal, M. T. (2013). Document Similarity Detection Using Semantic Social Network Analysis on RDF Citation Graph. *IEEE*.
- Mardeean, T., & Hidayah, I. (2016). The Comparation of Distance-Based Similarity Measure to Detection of Plagiarism in Indonesian Text. *Communication in Computer and Information Science*. Yogyakarta.
- Nugroho, E. (2011). *Perancangan Sistem Deteksi Plagiarisme Dokumen Teks Dengan Menggunakan Algoritma Rabin-Karp*. Malang: Universitas Brawijaya .
- Nurdiana, O., Jumadi, & Nursantika, D. (2016). Perbandingan Metode Cosine Similarity dengan Metode Jaccard Similarity Pada Aplikasi Pencarian Terjemah Al-Qur'an Dalam Bahasa Indonesia. *JOIN*.
- Pradiansyah, A. (2015). Makna plagiarisme bagi mahasiswa universitas muhammadiyah surakarta. *naskah publikasi universitas muhammadiyah surakarta*.
- Putra, D. A., Sujaini, H., & Pratiwi, H. S. (2015). Implementasi Algoritma Rabin-Karp untuk Membantu Pendeteksian Plagiat pada Karya Ilmiah. *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- Putra, D. A., Sujaini, H., & Pratiwi, H. S. (2015). Implementasi Algoritma Rabin-Karp untuk Membantu Pendeteksian Plagiat Pada Karya Ilmiah. *Jurnal Sistem dan Teknologi Informasi*.
- Rani, M., & Kaur, H. (2018). Similarity detection in sanskrit documents using binary search tree. *international journal of innovation & advancement in computer science*.
- Saptono, R. (2018). Combination of Cosine Similarity Method and conditional probability for plagiarism detection in the thesis documents vector space model. *journal of telecommunication*.
- Shivaji, S. K., & S, P. (2015). plagiarism detection by using rabin-karp string matching algorithm together. *international journal of computer applications*.
- Sinaga, & Akbar, H. F. (2018). *Implementasi Algoritma Rabin Karp Pada Pendeteksian Pengulangan Kata Kerja dalam Penulisan Bahasa Inggris*. Medan: Repositori Institusi USU.
- Sinaga, D. D., & Hansun, S. (2018). Indonesian Text Document Similarity Detection System Using Rabin-Karp and Confix Stripping Algorithms. *International Journal of Innovative Computing, Information and Control*.
- Stein, B., & Eissen, S. M. (2005). Near Similarity Search and Pplagiarsm Analysis. *From data and information analysis to knowledge engineering*.
- Sugiarti, Y. (2013). *Analisis dan Perancangan UML (Unified Modelling Language) Generated VB. 6*. Yogyakarta: Graha Ilmu.
- Sukmana, A., Kusriani, & Sunyoto, A. (2018). Perbandingan Penggunaan Stemming Pada Deteksi Kemiripan Dokumen Menggunakan Metode Rabin-Karp dan Jaccard Similarity. *Seminar Nasional Teknologi Informasi dan Multimedia*.
- Wibowo, A. T., Sudarmadi, K. W., & Barmawi, A. M. (2013). Comparison Btween Fingerprint and Winnowing Algorithm to Detect Plagiarsm Fraud on Bhasa Indonesia Documents. *International Conference if Information and Communication Technology*.
- Wicaksono, Y. A., & Suyanto. (2012). Analisis dan Implementasi Algoritma Rabin-Karp dan Algoritma Stemming Nazief-Adriani Pada Sistem Pendeteksi Plagiat Dokumen Teks Berbahasa Indonesia. *Telkom University*.
- Yu, X., Chen, X., & Shi, J. (2017). Vector Based Privacy-Preserving Document Similarity with LSA. *IEEE International Conference on Communication Software and Networks*.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN A BUKTI WAWANCARA

LAMPIRAN A BUKTI WAWANCARA

1. Daftar Pertanyaan Wawancara

Hari/Tanggal : Kamis, 24 Oktober 2019
Narasumber : Fitra Kurnia, S.Kom, M.Kom
Jabatan : Dosen Teknik Informatika UIN SUSKA RIAU

Pertanyaan :

1. Menurut anda, apa arti dari kesamaan dokumen?
2. Apakah anda pernah menemukan kesamaan antar dokumen mahasiswa?
3. Pada bentuk tugas yang seperti apa ditemukannya kesamaan dokumen antar mahasiswa?
4. Bagaimana bentuk praktik kesamaan dokumen yang sering dilakukan mahasiswa?
 - a. *Copy-paste* secara keseluruhan
 - b. *Copy-paste* sebagian
 - c. *Copy-paste* keseluruhan dan memparafrase sebagian
5. Bagaimana cara anda untuk memeriksa kesamaan antar dokumen mahasiswa?
6. Berapa lama waktu yang dibutuhkan untuk memeriksa kesamaan dokumen antar mahasiswa?
7. Apakah ada syarat-syarat tertentu yang mendasarkan bahwa sebuah dokumen dapat dikatakan sama atau plagiat?

Jawaban:

1. Dokumen yang dikatakan sama yaitu dokumen yang isi nya sama dengan dokumen lain, baik itu secara keseluruhan, sebagian, maupun parafrase sebagian ataupun dokumen yang mempunyai persentase kesamaan 40% keatas.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

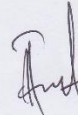
2. Pernah, kesamaan dokumen antar mahasiswa masih pernah ditemukan, tetapi dalam satu kelas hanya beberapa atau sebagian kecil mahasiswa yang dokumen tugas nya sama.

Masih terdapat rasa malas dalam diri sebagian mahasiswa yang membuat dirinya dalam mengerjakan tugas yang diberikan hanya menyalin tugas dari temannya sendiri, tidak ada persiapan atau kurang kepedulian terhadap tugas yang diberikan.

3. Tugas yang ditemukan bermacam-macam, tidak hanya pada makalah saja.
4. Semua bentuk praktik kesamaan dokumen pernah dilakukan, baik itu *copy-paste* secara keseluruhan, sebagian, maupun *copy-paste* keseluruhan namun kata-perkatanya diparafrase sebagian.
5. Dalam melakukan pemeriksaan dokumen tugas mahasiswa, belum menggunakan sistem atau aplikasi, pemeriksaan dokumen tugas dilakukan dengan membaca secara teliti satu persatu dokumen tugas mahasiswa.
6. Waktu yang dibutuhkan tergantung dengan seberapa banyak tugas yang diberikan, maksimal waktu yang dibutuhkan satu hari.

Dosen Teknik Informatika

UIN SUSKA RIAU,



(Fitra Kurnia, S.Kom, M.Pd)

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Daftar Pertanyaan Wawancara

Hari/Tanggal : Jum'at, 25 Oktober 2019
Narasumber : Sonya Meitarice, ST, M.Eng
Jabatan : Dosen Teknik Informatika UIN SUSKA RIAU

Pertanyaan :

1. Menurut anda, apa arti dari kesamaan dokumen?
2. Apakah anda pernah menemukan kesamaan antar dokumen mahasiswa?
3. Pada bentuk tugas yang seperti apa ditemukannya kesamaan dokumen antar mahasiswa?
4. Bagaimana bentuk praktik kesamaan dokumen yang sering dilakukan mahasiswa?
 - a. *Copy-paste* secara keseluruhan
 - b. *Copy-paste* sebagian
 - c. *Copy-paste* keseluruhan dan memparafrase sebagian
5. Bagaimana cara anda untuk memeriksa kesamaan antar dokumen mahasiswa?
6. Berapa lama waktu yang dibutuhkan untuk memeriksa kesamaan dokumen antar mahasiswa?
7. Apakah ada syarat-syarat tertentu yang mendasarkan bahwa sebuah dokumen dapat dikatakan sama atau plagiat?

Jawaban:

1. Kesamaan dokumen merupakan dokumen yang memiliki kesamaan lebih dari 30%.
2. Kesamaan dokumen antar mahasiswa masih ada.
3. Bentuk tugas yang ditemukan kesamaan beragam, ada makalah, tugas-tugas menghitung, dan lain sebagainya.
4. Bentuk praktik kesamaan dokumen yang terjadi semuanya ada. Ada Mahasiswa yang *copy-paste* tugas mahasiswa lain secara keseluruhan, mahasiswa yang *copy-paste* sebagian, ada juga mahasiswa yang

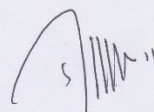
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

menegrikan tugas menyalin keseluruhan namun diparafrase sebagian, hingga mahasiswa yang melakukan *copy-paste* dengan merubah tata letak kata dari dokumen aslinya.

5. Cara untuk memeriksa kesamaan antar dokumen tugas mahasiswa, dapat dilihat dari penyusunan bab, dari jenis dan ukuran *font* yang dipilih, tata letak kalimat, dan lain sebagainya.
6. Waktu yang dibutuhkan tidak cukup lama untuk memeriksa dokumen tugas mahasiswa.
7. Syarat yang digunakan tergantung pada SOP universitas, standar kesamaan yang digunakan yaitu dibawah 30%, dokumen yang memiliki kesamaan diatas 30-35% dianggap plagiat.

Dosen Teknik Informatika
UIN SUSKA RIAU,



(Sonya Meitarice, ST, M.Eng)

A-4

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. Daftar Pertanyaan Wawancara

Hari/Tanggal : Senin, 28 Oktober 2019
Narasumber : Eka Pandu Chyntia, ST, M.Kom
Jabatan : Dosen Teknik Informatika UIN SUSKA RIAU

Pertanyaan :

1. Menurut anda, apa arti dari kesamaan dokumen?
2. Apakah anda pernah menemukan kesamaan antar dokumen mahasiswa?
3. Pada bentuk tugas yang seperti apa ditemukannya kesamaan dokumen antar mahasiswa?
4. Bagaimana bentuk praktik kesamaan dokumen yang sering dilakukan mahasiswa?
 - a. *Copy-paste* secara keseluruhan
 - b. *Copy-paste* sebagian
 - c. *Copy-paste* keseluruhan dan memparafrase sebagian
5. Bagaimana cara anda untuk memeriksa kesamaan antar dokumen mahasiswa?
6. Berapa lama waktu yang dibutuhkan untuk memeriksa kesamaan dokumen antar mahasiswa?
7. Apakah ada syarat-syarat tertentu yang mendasarkan bahwa sebuah dokumen dapat dikatakan sama atau plagiat?

Jawaban:

1. Kesamaan dokumen memiliki dua arti, yang pertama dokumen dengan isi sama persis dengan dokumen lain, dan yang kedua dokumen yang memiliki bahasa atau kata yang hanya dibolak-balik namun dengan inti yang sama dari dokumen lain.
2. Dalam memberikan sebuah tugas kepada mahasiswa terdapat beberapa *rule* diantaranya:
 - Tugas yang diberikan tidak akan pernah dengan soal-an yang sama antara mahasiswa satu dengan yang lainnya.

A-5

Hak Cipta Dilindungi Undang-Undang

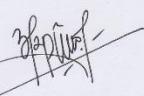
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- Jika ada yang ketahuan sama maka akan diberi nilai 0

Namun dari dua *rule* yang sudah diberikan tersebut, masih juga pernah bahkan sering dijumpai kesamaan dari tugas yang diberikan yang dilakukan oleh mahasiswa yang tidak mendengar atau tidak memperhatikan *rule* yang diberikan.

3. Bentuk tugas yang diberikan beragam, seperti makalah, *paper* dan terdapat juga tugas yang berbentuk pendapat mengenai topik tertentu, dan sebagainya. Pada tugas yang diberikan jika hanya sebanyak satu lembar maka dikerjakan dengan tulis tangan, tetapi jika memerlukan dua sampai tiga lembar, dikerjakan dengan diketik.
4. Bentuk kesamaan bermacam-macam, dan semua nya ada. Baik *copy-paste* keseluruhan, sebagian, maupun parafrase sebagian.
5. Pengecekan dokumen dilakukan dengan menggunakan *tools plagiarism checker*.
6. Karena aplikasi yang digunakan secara *online*, maka waktu yang dibutuhkan tergantung pada koneksi internet.
7. Syarat-syaratnya tergantung bobot dari soal, jika tugas membutuhkan banyak referensi, maka toleransi kesamaan akan dinaikkan persentasenya. Contohnya pada tugas makalah yang harus menggunakan referensi sekian banyak, maka akan diberi aturan, jika ditemukan pada aplikasi yang digunakan untuk pemeriksaan kesamaan misalnya dengan batas diatas 30% maka akan diberi nilai 0.

Dosen Teknik Informatika
UIN SUSKA RIAU,


(Eka Pandu Chyathia, ST, M.Kom)
Cynthia.

A-6



LAMPIRAN B

TABEL HASIL PERHITUNGAN NILAI HASH MENGGUNAKAN MS. OFFICE EXCEL

1. Perhitungan Nilai *Hash* Dokumen 1

STRING												MOD	HASH	
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
b	e	r	a	s	a	l						191692400	100007	78988
98	101	114	97	115	97	108								
d	a	r	i									146196	100007	46189
100	97	114	105											
k	a	t	a									155527	100007	55520
107	97	116	97											
m	o	t	i	f								1758903	100007	58784
109	111	116	105	102										
y	a	n	g									174101	100007	74094
121	97	110	103											
b	e	r	a	r	t	i						191692485	100007	79073
98	101	114	97	114	116	105								
d	a	y	a									146265	100007	46258
100	97	121	97											
u	p	a	y	a								1875234	100007	75108



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
117	112	97	121	97										
y	a	n	g									174101	100007	74094
121	97	110	103											
m	e	n	d	o	r	o	n	g				25546090739	100007	2638
109	101	110	100	111	114	111	110	103						
s	e	s	e	o	r	a	n	g				26841261187	100007	82436
115	101	115	101	111	114	97	110	103						
u	n	t	u	k								1874837	100007	74711
117	110	116	117	107										
m	e	l	a	k	u	k	a	n				25542009273	100007	21459
109	101	108	97	107	117	107	97	110						
s	e	s	u	a	t	u						221848238	100007	32712
115	101	115	117	97	116	117								
d	o	r	o	n	g	a	n					2165505386	100007	53815
100	111	114	111	110	103	97	110							
d	a	s	a	r								1608303	100007	8191
100	97	115	97	114										
y	a	n	g									174101	100007	74094
121	97	110	103											
m	e	n	g	g	e	r	a	k	k	a	n	34002311234859	100007	39675
109	101	110	103	103	101	114	97	107	107	97	110			
t	i	n	g	k	a	h						224173152	100007	57465
116	105	110	103	107	97	104								



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
l	a	k	u									156779	100007	56772
108	97	107	117											
m	a	n	u	s	i	a						210503500	100007	88772
109	97	110	117	115	105	97								
d	i	s	e	b	u	t						195897862	100007	84156
100	105	115	101	98	117	116								
s	e	b	a	g	a	i						221573226	100007	57721
115	101	98	97	103	97	105								
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
a	d	a	l	a	h							17229393	100007	28189
97	100	97	108	97	104									
k	e	k	u	a	t	a	n					2283144732	100007	84929
107	101	107	117	97	116	97	110							
y	a	n	g									174101	100007	74094
121	97	110	103											
m	u	n	c	u	l							19427340	100007	25982
109	117	110	99	117	108									
d	a	r	i									146196	100007	46189
100	97	114	105											
d	a	l	a	m								1607451	100007	7339



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
100	97	108	97	109										
d	i	r	i									147164	100007	47157
100	105	114	105											
s	e	s	e	o	r	a	n	g				26841261187	100007	82436
115	101	115	101	111	114	97	110	103						
u	n	t	u	k								1874837	100007	74711
117	110	116	117	107										
m	e	n	d	a	p	a	t	k	a	n		3091051660136	100007	1665
109	101	110	100	97	112	97	116	107	97	110				
t	u	j	u	a	n							20551333	100007	49898
116	117	106	117	97	110									
y	a	n	g									174101	100007	74094
121	97	110	103											
d	i	i	n	g	i	n	k	a	n			260561762362	100007	24317
100	105	105	110	103	105	110	107	97	110					
d	a	n										13277	100007	13277
100	97	110												
b	e	r	u	s	a	h	a					2108909273	100007	61664
98	101	114	117	115	97	104	97							
m	e	l	a	k	u	k	a	n				25542009273	100007	21459
109	101	108	97	107	117	107	97	110						
p	e	r	u	b	a	h	a	n				26198777550	100007	43767
112	101	114	117	98	97	104	97	110						



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
t	i	n	g	k	a	h						224173152	100007	57465
116	105	110	103	107	97	104								
l	a	k	u									156779	100007	56772
108	97	107	117											
y	a	n	g									174101	100007	74094
121	97	110	103											
l	e	b	i	h								1728776	100007	28657
108	101	98	105	104										
b	a	i	k									143437	100007	43430
98	97	105	107											
d	e	m	i									146625	100007	46618
100	101	109	105											
m	e	m	e	n	u	h	i					2322224416	100007	61876
109	101	109	101	110	117	104	105							
k	e	i	n	g	i	n	a	n	n	y	a	33421405113996	100007	79397
107	101	105	110	103	105	110	97	110	110	121	97			
s	e	m	a	k	i	n						221734854	100007	19335
115	101	109	97	107	105	110								
b	e	s	a	r								1584345	100007	84240
98	101	115	97	114										
d	o	r	o	n	g	a	n					2165505386	100007	53815
100	111	114	111	110	103	97	110							
d	a	r	i									146196	100007	46189



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
100	97	114	105											
d	a	l	a	m								1607451	100007	7339
100	97	108	97	109										
d	i	r	i									147164	100007	47157
100	105	114	105											
s	e	s	e	o	r	a	n	g				26841261187	100007	82436
115	101	115	101	111	114	97	110	103						
m	a	k	a									158090	100007	58083
109	97	107	97											
s	e	m	a	k	i	n						221734854	100007	19335
115	101	109	97	107	105	110								
b	e	s	a	r								1584345	100007	84240
98	101	115	97	114										
p	u	l	a									164514	100007	64507
112	117	108	97											
u	s	a	h	a								1879040	100007	78914
117	115	97	104	97										
y	a	n	g									174101	100007	74094
121	97	110	103											
d	i	l	a	k	u	k	a	n	n	y	a	31532359020006	100007	9373
100	105	108	97	107	117	107	97	110	110	121	97			
u	n	t	u	k								1874837	100007	74711
117	110	116	117	107										



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING											MOD	HASH	
m	e	n	c	a	p	a	i				2322338200	100007	75653
109	101	110	99	97	112	97	105						
t	u	j	u	a	n						20551333	100007	49898
116	117	106	117	97	110								
m	o	t	i	v	a	s	i				2341134296	100007	70433
109	111	116	105	118	97	115	105						
s	a	n	g	a	t						20101098	100007	99698
115	97	110	103	97	116								
d	i	p	e	r	l	u	k	a	n		260684053311	100007	6705
100	105	112	101	114	108	117	107	97	110				
d	a	l	a	m							1607451	100007	7339
100	97	108	97	109									
p	r	o	s	e	s						19869668	100007	68282
112	114	111	115	101	115								
b	e	l	a	j	a	r					191603471	100007	90066
98	101	108	97	106	97	114							
s	e	b	a	b							1831169	100007	31043
115	101	98	97	98									
s	e	s	e	o	r	a	n	g			26841261187	100007	82436
115	101	115	101	111	114	97	110	103					
t	i	d	a	k							1851385	100007	51259
116	105	100	97	107									
a	k	a	n								143231	100007	43224



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
97	107	97	110											
m	e	l	a	k	u	k	a	n				25542009273	100007	21459
109	101	108	97	107	117	107	97	110						
a	k	t	i	v	i	t	a	s				23102232796	100007	15754
97	107	116	105	118	105	116	97	115						
b	e	l	a	j	a	r						191603471	100007	90066
98	101	108	97	106	97	114								
j	i	k	a									155065	100007	55058
106	105	107	97											
t	i	d	a	k								1851385	100007	51259
116	105	100	97	107										
m	e	m	p	u	n	y	a	i				25546335381	100007	47266
109	101	109	112	117	110	121	97	105						
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
d	a	l	a	m								1607451	100007	7339
100	97	108	97	109										
b	e	l	a	j	a	r						191603471	100007	90066
98	101	108	97	106	97	114								
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
m	e	n	i	m	b	u	l	k	a	n		3091168382181	100007	15541
109	101	110	105	109	98	117	108	107	97	110				



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
g	a	i	r	a	h							18163150	100007	61883
103	97	105	114	97	104									
s	e	m	a	n	g	a	t					2439087118	100007	16395
115	101	109	97	110	103	97	116							
d	a	n										13277	100007	13277
100	97	110												
m	e	r	a	s	a							19198133	100007	96796
109	101	114	97	115	97									
s	e	n	a	n	g							20159066	100007	57659
115	101	110	97	110	103									
u	n	t	u	k								1874837	100007	74711
117	110	116	117	107										
b	e	l	a	j	a	r						191603471	100007	90066
98	101	108	97	106	97	114								
s	e	m	a	k	i	n						221734854	100007	19335
115	101	109	97	107	105	110								
b	e	s	a	r								1584345	100007	84240
98	101	115	97	114										
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
s	e	s	e	o	r	a	n	g				26841261187	100007	82436
115	101	115	101	111	114	97	110	103						
m	a	k	a									158090	100007	58083



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH		
109	97	107	97												
s	e	m	a	k	i	n							221734854	100007	19335
115	101	109	97	107	105	110									
b	e	s	a	r									1584345	100007	84240
98	101	115	97	114											
p	u	l	a										164514	100007	64507
112	117	108	97												
e	n	e	r	g	i								18026124	100007	24864
101	110	101	114	103	105										
y	a	n	g										174101	100007	74094
121	97	110	103												
d	i	m	i	l	i	k	i	n	y	a			2866948198662	100007	26337
100	105	109	105	108	105	107	105	110	121	97					
u	n	t	u	k									1874837	100007	74711
117	110	116	117	107											
b	e	l	a	j	a	r							191603471	100007	90066
98	101	108	97	106	97	114									



Hasil perhitungan dengan $K\text{-gram} = 4$

No	HASIL K-GRAM = 4		No	HASIL K-GRAM = 4	
1	motivasi berasal dari kata	251130	54	dorongan dari dalam diri	154500
2	berasal dari kata motif	239481	55	dari dalam diri seseorang	183121
3	dari kata motif yang	234587	56	dalam diri seseorang maka	195015
4	kata motif yang berarti	267471	57	diri seseorang maka semakin	207011
5	motif yang berarti daya	258209	58	seseorang maka semakin besar	244094
6	yang berarti daya upaya	274533	59	maka semakin besar pula	226165
7	berarti daya upaya yang	274533	60	semakin besar pula usaha	246996
8	daya upaya yang mendorong	198098	61	besar pula usaha yang	301755
9	upaya yang mendorong seseorang	234276	62	pula usaha yang dilakukannya	226888
10	yang mendorong seseorang untuk	233879	63	usaha yang dilakukannya untuk	237092
11	mendorong seseorang untuk melakukan	181244	64	yang dilakukannya untuk mencapai	233831
12	seseorang untuk melakukan sesuatu	211318	65	dilakukannya untuk mencapai tujuan	209635
13	untuk melakukan sesuatu dorongan	182697	66	untuk mencapai tujuan motivasi	270695
14	melakukan sesuatu dorongan dasar	116177	67	mencapai tujuan motivasi sangat	295682
15	sesuatu dorongan dasar yang	168812	68	tujuan motivasi sangat diperlukan	226734
16	dorongan dasar yang menggerakkan	175775	69	motivasi sangat diperlukan dalam	184175
17	dasar yang menggerakkan tingkah	179425	70	sangat diperlukan dalam proses	182024
18	yang menggerakkan tingkah laku	228006	71	diperlukan dalam proses belajar	172392
19	menggerakkan tingkah laku manusia	242684	72	dalam proses belajar sebab	196730
20	tingkah laku manusia disebut	287165	73	proses belajar sebab seseorang	271827
21	laku manusia disebut sebagai	287421	74	belajar sebab seseorang tidak	254804



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

No	HASIL K-GRAM = 4		No	HASIL K-GRAM = 4	
22	manusia disebut sebagai motivasi	301082	75	sebab seseorang tidak akan	207962
23	disebut sebagai motivasi motivasi	282743	76	seseorang tidak akan melakukan	198378
24	sebagai motivasi motivasi adalah	226776	77	tidak akan melakukan aktivitas	131696
25	motivasi motivasi adalah kekuatan	253984	78	akan melakukan aktivitas belajar	170503
26	motivasi adalah kekuatan yang	257645	79	melakukan aktivitas belajar jika	182337
27	adalah kekuatan yang muncul	213194	80	aktivitas belajar jika tidak	212137
28	kekuatan yang muncul dari	231194	81	belajar jika tidak mempunyai	243649
29	yang muncul dari dalam	153604	82	jika tidak mempunyai motivasi	224016
30	muncul dari dalam diri	126667	83	tidak mempunyai motivasi dalam	176297
31	dari dalam diri seseorang	183121	84	mempunyai motivasi dalam belajar	215104
32	dalam diri seseorang untuk	211643	85	motivasi dalam belajar motivasi	238271
33	diri seseorang untuk mendapatkan	205969	86	dalam belajar motivasi menimbulkan	183379
34	seseorang untuk mendapatkan tujuan	208710	87	belajar motivasi menimbulkan gairah	237923
35	untuk mendapatkan tujuan yang	200368	88	motivasi menimbulkan gairah semangat	164252
36	mendapatkan tujuan yang diinginkan	149974	89	menimbulkan gairah semangat dan	107096
37	tujuan yang diinginkan dan	161586	90	gairah semangat dan merasa	188351
38	yang diinginkan dan berusaha	173352	91	semangat dan merasa senang	184127
39	diinginkan dan berusaha melakukan	120717	92	dan merasa senang untuk	242443
40	dan berusaha melakukan perubahan	140167	93	merasa senang untuk belajar	319232
41	berusaha melakukan perubahan tingkah	184355	94	senang untuk belajar semakin	241771
42	melakukan perubahan tingkah laku	179463	95	untuk belajar semakin besar	268352
43	perubahan tingkah laku yang	232098	96	belajar semakin besar motivasi	264074
44	tingkah laku yang lebih	216988	97	semakin besar motivasi seseorang	256444



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

No	HASIL K-GRAM = 4		No	HASIL K-GRAM = 4	
45	laku yang lebih baik	202953	98	besar motivasi seseorang maka	295192
46	yang lebih baik demi	192799	99	motivasi seseorang maka semakin	230287
47	lebih baik demi memenuhi	180581	100	seseorang maka semakin besar	244094
48	baik demi memenuhi keinginannya	231321	101	maka semakin besar pula	226165
49	demi memenuhi keinginannya semakin	207226	102	semakin besar pula energi	192946
50	memenuhi keinginannya semakin besar	244848	103	besar pula energi yang	247705
51	keinginannya semakin besar dorongan	236787	104	pula energi yang dimilikinya	189802
52	semakin besar dorongan dari	203579	105	energi yang dimilikinya untuk	200006
53	besar dorongan dari dalam	191583	106	yang dimilikinya untuk belajar	265208

2. Perhitungan Nilai Hash Dokumen 2

STRING												MOD	HASH	
p	e	n	g	e	r	t	i	a	n			288084551469	100007	86968
112	101	110	103	101	114	116	105	97	110					
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
m	e	n	u	r	u	t						211147734	100007	32957
109	101	110	117	114	117	116								
p	a	r	a									162160	100007	62153
112	97	114	97											
a	h	l	i									142984	100007	42977



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
97	104	108	105											
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
b	e	r	a	s	a	l						191692400	100007	78988
98	101	114	97	115	97	108								
d	a	r	i									146196	100007	46189
100	97	114	105											
k	a	t	a									155527	100007	55520
107	97	116	97											
m	o	t	i	f								1758903	100007	58784
109	111	116	105	102										
y	a	n	g									174101	100007	74094
121	97	110	103											
d	i	a	r	t	i	k	a	n				23674094521	100007	37453
100	105	97	114	116	105	107	97	110						
s	e	b	a	g	a	i						221573226	100007	57721
115	101	98	97	103	97	105								
d	a	y	a									146265	100007	46258
100	97	121	97											
u	p	a	y	a								1875234	100007	75108
117	112	97	121	97										
y	a	n	g									174101	100007	74094



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
121	97	110	103											
m	e	n	d	o	r	o	n	g				25546090739	100007	2638
109	101	110	100	111	114	111	110	103						
s	e	s	e	o	r	a	n	g				26841261187	100007	82436
115	101	115	101	111	114	97	110	103						
u	n	t	u	k								1874837	100007	74711
117	110	116	117	107										
m	e	l	a	k	u	k	a	n				25542009273	100007	21459
109	101	108	97	107	117	107	97	110						
s	e	s	u	a	t	u						221848238	100007	32712
115	101	115	117	97	116	117								
m	e	n	u	r	u	t						211147734	100007	32957
109	101	110	117	114	117	116								
s	a	r	d	i	m	a	n					2432844117	100007	73835
115	97	114	100	105	109	97	110							
2	0	0	6	7								802395	100007	2339
50	48	48	54	55										
m	o	t	i	f								1758903	100007	58784
109	111	116	105	102										
m	e	r	u	p	a	k	a	n				25555906244	100007	17457
109	101	114	117	112	97	107	97	110						
d	a	y	a									146265	100007	46258



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
100	97	121	97											
p	e	n	g	g	e	r	a	k				26189516328	100007	83196
112	101	110	103	103	101	114	97	107						
d	a	r	i									146192	100007	46185
100	97	114	101											
d	a	l	a	m								1607451	100007	7339
100	97	108	97	109										
u	n	t	u	k								1874837	100007	74711
117	110	116	117	107										
m	e	l	a	k	u	k	a	n				25542009273	100007	21459
109	101	108	97	107	117	107	97	110						
k	e	g	i	a	t	a	n					2282324836	100007	65089
107	101	103	105	97	116	97	110							
u	n	t	u	k								1874837	100007	74711
117	110	116	117	107										
m	e	n	c	a	p	a	i					2322338200	100007	75653
109	101	110	99	97	112	97	105							
t	u	j	u	a	n							20551333	100007	49898
116	117	106	117	97	110									
d	a	l	a	m								1607451	100007	7339
100	97	108	97	109										
p	r	o	s	e	s							19869668	100007	68282



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH		
112	114	111	115	101	115										
b	e	l	a	j	a	r							191603471	100007	90066
98	101	108	97	106	97	114									
m	o	t	i	v	a	s	i						2341134296	100007	70433
109	111	116	105	118	97	115	105								
s	a	n	g	a	t								20101098	100007	99698
115	97	110	103	97	116										
d	i	p	e	r	l	u	k	a	n				260684053311	100007	6705
100	105	112	101	114	108	117	107	97	110						
s	e	b	a	b									1831169	100007	31043
115	101	98	97	98											
s	e	s	e	o	r	a	n	g					26841261187	100007	82436
115	101	115	101	111	114	97	110	103							
y	a	n	g										174101	100007	74094
121	97	110	103												
t	i	d	a	k									1851385	100007	51259
116	105	100	97	107											
m	e	m	p	u	n	y	a	i					25546335381	100007	47266
109	101	109	112	117	110	121	97	105							
m	o	t	i	v	a	s	i						2341134296	100007	70433
109	111	116	105	118	97	115	105								
d	a	l	a	m									1607451	100007	7339
100	97	108	97	109											



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
b	e	l	a	j	a	r						191603471	100007	90066
98	101	108	97	106	97	114								
t	i	d	a	k								1851385	100007	51259
116	105	100	97	107										
a	k	a	n									143121	100007	43114
97	107	97												
m	u	n	g	k	i	n						213704931	100007	89979
109	117	110	103	107	105	110								
m	e	l	a	k	s	a	n	a	k	a	n	33996409375885	100007	93806
109	101	108	97	107	115	97	110	97	107	97	110			
a	k	t	i	v	i	t	a	s				23102232796	100007	15754
97	107	116	105	118	105	116	97	115						
b	e	l	a	j	a	r						191603471	100007	90066
98	101	108	97	106	97	114								
m	o	t	i	v	a	s	i					2341134296	100007	70433
109	111	116	105	118	97	115	105							
d	i	p	e	r	l	u	k	a	n			260684053311	100007	6705
100	105	112	101	114	108	117	107	97	110					
d	a	l	a	m								1607451	100007	7339
100	97	108	97	109										
m	e	n	e	n	t	u	k	a	n			281008645588	100007	76365
109	101	110	101	110	116	117	107	97	110					
i	n	t	e	n	s	i	t	a	s			273622978260	100007	25994



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

STRING												MOD	HASH	
105	110	116	101	110	115	105	116	97	115					
u	s	a	h	a								1879040	100007	78914
117	115	97	104	97										
b	e	l	a	j	a	r						191603471	100007	90066
98	101	108	97	106	97	114								
b	a	g	i									143413	100007	43406
98	97	103	105											
p	a	r	a									162160	100007	62153
112	97	114	97											
s	i	s	w	a								1838791	100007	38665
115	105	115	119	97										

Hasil perhitungan dengan $K\text{-gram} = 4$

No	HASIL K-GRAM = 4		No	HASIL K-GRAM = 4	
1	pengertian motivasi menurut para	252511	33	kegiatan untuk mencapai tujuan	265351
2	motivasi menurut para ahli	208520	34	untuk mencapai tujuan dalam	207601
3	menurut para ahli motivasi	208520	35	mencapai tujuan dalam proses	201172
4	para ahli motivasi berasal	254551	36	tujuan dalam proses belajar	215585
5	ahli motivasi berasal dari	238587	37	dalam proses belajar motivasi	236120
6	motivasi berasal dari kata	251130	38	proses belajar motivasi sangat	328479
7	berasal dari kata motif	239481	39	belajar motivasi sangat diperlukan	266902



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

No	HASIL K-GRAM = 4		No	HASIL K-GRAM = 4	
8	dari kata motif yang	234587	40	motivasi sangat diperlukan sebab	207879
9	kata motif yang diartikan	225851	41	sangat diperlukan sebab seseorang	219882
10	motif yang diartikan sebagai	228052	42	diperlukan sebab seseorang yang	194278
11	yang diartikan sebagai daya	215526	43	sebab seseorang yang tidak	238832
12	diartikan sebagai daya upaya	216540	44	seseorang yang tidak mempunyai	255055
13	sebagai daya upaya yang	253181	45	yang tidak mempunyai motivasi	243052
14	daya upaya yang mendorong	198098	46	tidak mempunyai motivasi dalam	176297
15	upaya yang mendorong seseorang	234276	47	mempunyai motivasi dalam belajar	215104
16	yang mendorong seseorang untuk	233879	48	motivasi dalam belajar tidak	219097
17	mendorong seseorang untuk melakukan	181244	49	dalam belajar tidak akan	191888
18	seseorang untuk melakukan sesuatu	211318	50	belajar tidak akan mungkin	274528
19	untuk melakukan sesuatu menurut	161839	51	tidak akan mungkin melaksanakan	278268
20	melakukan sesuatu menurut sardiman	160963	52	akan mungkin melaksanakan aktivitas	242763
21	sesuatu menurut sardiman 200673	165284	53	mungkin melaksanakan aktivitas belajar	289605
22	menurut sardiman 200673 motif	191356	54	melaksanakan aktivitas belajar motivasi	270059
23	sardiman 200673 motif merupakan	175856	55	aktivitas belajar motivasi diperlukan	182958
24	200673 motif merupakan daya	148279	56	belajar motivasi diperlukan dalam	174543
25	motif merupakan daya penggerak	205695	57	motivasi diperlukan dalam menentukan	160842
26	merupakan daya penggerak dari	193100	58	diperlukan dalam menentukan intensitas	116403
27	daya penggerak dari dalam	182982	59	dalam menentukan intensitas usaha	188612
28	penggerak dari dalam untuk	211435	60	menentukan intensitas usaha belajar	271339
29	dari dalam untuk melakukan	149698	61	intensitas usaha belajar bagi	238380



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan
- Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

No	HASIL K-GRAM = 4		No	HASIL K-GRAM = 4	
30	dalam untuk melakukan kegiatan	168598	62	usaha belajar bagi para	274539
31	untuk melakukan kegiatan untuk	235970	63	belajar bagi para siswa	234290
32	melakukan kegiatan untuk mencapai	236912			



DAFTAR RIWAYAT HIDUP

Informasi Personal



Nama : Elvy Marisha
Tempat Tanggal Lahir : Duri, 12 Maret 1997
Jenis Kelamin : Perempuan
Tinggi Badan : 156 cm
Berat Badan : 46 kg
Anak Ke : 2 dari 3 Bersaudara
Kebangsaan : Indonesia
Agama : Islam

Alamat

Sekarang : Jl. HR. Soebrantas, Perumahan Villa Pesona Panam, Blok K-15, Pekanbaru
No. HP : 0853-6314-2080
Email : elvy.marisha@students.uin-suska.ac.id

Riwayat Pendidikan

Tahun 2004– 2009 : SD Negeri Centre 022 Mandau
Tahun 2009 – 2012 : SMP Negeri 04 Mandau
Tahun 2012 – 2015 : SMA Negeri 01 Mandau
Tahun 2015 – 2019 : S1 Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.